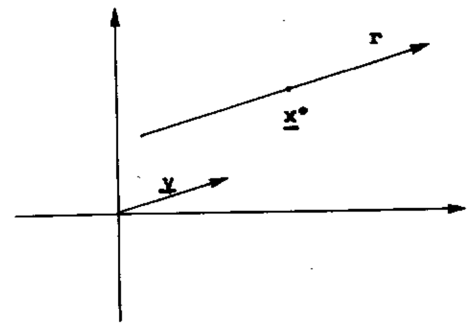
Capitolo II

CALCOLO DIFFERENZIALE PER LE FUNZIONI DI N VARIABILI

1. Derivata in una direzione.

Consideriamo lo spazio cartesiano R^n $(n\geq 1)$. Denotiamo (°) con $\underline{x}$ il punto generico di R^n, con $x_1,\dots,x_n$ le sue coordinate e con $\underline{e}_1,\dots,\underline{e}_n$ i versori degli assi coordinati.
Sia $\underline{f}$ una funzione definita in un aperto A di R^n a valori in R^m $(m\geq 1)$: $\underline{f}=(f_1,\dots,f_m)$.
Sia $\underline{x}^\circ=(x_1^\circ,\dots,x_n^\circ)$ un punto di A e $\underline{v}=(v_1,\dots,v_n)$ un versore assegnato di R^n. Diciamo r la retta (orientata!) passante per $\underline{x}^\circ$ e avente direzione e verso di $\underline{v}$ e consideriamo la restrizione di $\underline{f}$ alla retta r, cioè la funzione $\underline{\varphi}:R\to R^m$ della variabile reale t definita da



$$\underline{\varphi}(t) = \underline{f}(\underline{x}^\circ+t\underline{v}) \quad .$$

Se la funzione $\underline{\varphi}$ è derivabile rispetto a t nel punto $t=0$ con derivata finita, diremo che <u>esiste nel punto $\underline{x}^\circ$ la derivata</u> (parziale prima) <u>di $\underline{f}$ nella direzione di $\underline{v}$</u> (o <u>secondo il versore</u> $\underline{v}$) $D_{\underline{v}}\underline{f}(\underline{x}^\circ)$ assegnata da

$$(1.1) \qquad D_{\underline{v}}\underline{f}(\underline{x}^\circ) = \frac{d\underline{\varphi}}{dt}(0) = \lim_{t\to 0}\frac{\underline{f}(\underline{x}^\circ+t\underline{v})-\underline{f}(\underline{x}^\circ)}{t} \quad .$$

Ad esempio, consideriamo la funzione $f(x_1,x_2)=x_1^2 \operatorname{sen} x_2$, $(x_1,x_2\in R)$ e sia $\underline{x}^\circ=(1,\pi)$, $\underline{v}=(\frac{1}{2},\frac{\sqrt{3}}{2})$. Risulta

(°) In questo capitolo e nel successivo Cap. IV.

$$f(\underline{x}^\circ+t\underline{v}) = f(1+\tfrac{t}{2}, \pi+\tfrac{t\sqrt{3}}{2}) = (1+\tfrac{t}{2})^2 \operatorname{sen}(\pi+\tfrac{t\sqrt{3}}{2})$$

e quindi

$$D_{\underline{v}} f(\underline{x}^\circ) = -\frac{\sqrt{3}}{2} .$$

Se $n=1$ e $\underline{v}$ è il versore dell'asse delle ascisse, si ricade nella ben nota definizione di derivata prima delle funzioni di una sola variabile reale.
Se $n>1$, la derivata secondo il versore $\underline{e}_i$ dell'i-mo asse cartesiano nel punto $\underline{x}^\circ$ viene chiamata <u>derivata parziale prima di $\underline{f}$ rispetto a</u> x_i e indicata col simbolo $\frac{\partial f}{\partial x_i}(\underline{x}^\circ)$.

Da (1.1) si ricava

$$\frac{\partial f}{\partial x_i}(\underline{x}^\circ) = D_{\underline{e}_i}\underline{f}(\underline{x}^\circ) = \lim_{t\to 0}\frac{\underline{f}(\underline{x}^\circ+t\underline{e}_i)-\underline{f}(\underline{x}^\circ)}{t}$$

$$= \lim_{t\to 0}\frac{\underline{f}(x_1^\circ,\dots,x_{i-1}^\circ,x_i^\circ+t,x_{i+1}^\circ,\dots,x_n^\circ)-\underline{f}(x_1^\circ,\dots,x_i^\circ,\dots,x_n^\circ)}{t} .$$

Di conseguenza risulta :

$$(1.2)\qquad \frac{\partial f}{\partial x_i}(\underline{x}^\circ) = \left(\frac{d\, \underline{f}(x_1^\circ,\dots,x_{i-1}^\circ,x_i,x_{i+1}^\circ,\dots,x_n^\circ)}{d\, x_i}\right)_{x_i=x_i^\circ}$$

Si consideri, ad esempio, la funzione $f: R^2 \to R$ così definita:

$$f(x_1,x_2) = \begin{cases} \dfrac{x_1}{x_2} & \text{se } x_2\neq 0 \\ 0 & \text{se } x_2=0 \end{cases} .$$

Sia $\underline{x}^\circ=(x_1^\circ,x_2^\circ)$ con $x_2^\circ\neq 0$. Allora in tale punto esistono entrambe le derivate parziali prime e risulta

$$\frac{\partial f}{\partial x_1}(\underline{x}^\circ) = \left(\frac{d}{dx_1}\frac{x_1}{x_2^\circ}\right)_{x_1=x_1^\circ} = \frac{1}{x_2^\circ}$$

$$\frac{\partial f}{\partial x_2}(\underline{x}^\circ) = \left(\frac{d}{dx_2}\frac{x_1^\circ}{x_2}\right)_{x_2=x_2^\circ} = -\frac{x_1^\circ}{(x_2^\circ)^2} .$$

Sia $\underline{x}^\circ=(0,0)$. Anche in questo punto esistono entrambe le derivate parziali prime e sono entrambe nulle. Infatti le funzioni $f(x_1,0)$ e $f(0,x_2)$ sono identicamente nulle.
Infine sia $\underline{x}^\circ=(x_1^\circ,0)$ con $x_1^\circ\neq 0$. Allora in tale punto esiste $\frac{\partial f}{\partial x_1}$ ed è nulla, perchè $f(x_1,0)$ è identicamente nulla. Invece non esiste $\frac{\partial f}{\partial x_2}$; infatti risulta

$$f(x_1^\circ,x_2) = \begin{cases} \dfrac{x_1^\circ}{x_2} & \text{se } x_2\neq 0 \\ 0 & \text{se } x_2=0 \end{cases}$$

e tale funzione non è derivabile nel punto $x_2=0$.

La matrice $J(\underline{f};\underline{x}^\circ)=\left[\frac{\partial f_j}{\partial x_i}(\underline{x}^\circ)\right]_{\substack{j=1,\dots,m\\ i=1,\dots,n}}$ viene chiamata <u>matrice jacobiana</u> o <u>matrice derivata</u> di $\underline{f}$ in $\underline{x}^\circ$.
Se $n=m$, il suo determinante viene chiamato <u>determinante jacobiano</u> di $\underline{f}$ in $\underline{x}^\circ$ e sovente indicato col simbolo $J\begin{pmatrix} f_1,\dots,f_n \\ x_1,\dots,x_n\end{pmatrix}_{\underline{x}=\underline{x}^\circ}$ oppure con $\left(\frac{D(f_1,\dots,f_n)}{D(x_1,\dots,x_n)}\right)_{\underline{x}=\underline{x}^\circ}$

Ad esempio, la funzione $\underline{f}: R^2\to R^2$ definita da $\underline{f}(x_1,x_2)=(x_1x_2, \operatorname{sen} x_2)$ ha come matrice jacobiana nel generico punto $\underline{x}$ di R^2

$$\begin{bmatrix} x_2 & x_1 \\ 0 & \cos x_2 \end{bmatrix}$$

e il suo determinante jacobiano ha valore $x_2 \cos x_2$.

Se $m=n=1$ il determinante jacobiano coincide con la derivata prima $\frac{df}{dx}$.
Se $m=1$ e $n>1$, il vettore che ha come componenti le derivate parziali prime nel punto $\underline{x}^\circ$ viene sovente chiamato <u>gradiente</u> di f in $\underline{x}^\circ$. E' quindi, per definizione :

$$\operatorname{grad} f(\underline{x}^\circ) = \left(\frac{\partial f}{\partial x_1}(\underline{x}^\circ),\dots,\frac{\partial f}{\partial x_n}(\underline{x}^\circ)\right) .$$

Ad esempio, la funzione

$$f(x_1,x_2) = \begin{cases} \dfrac{x_1}{x_2} & \text{se } x_2\neq 0 \\ 0 & \text{se } x_2=0 \end{cases}$$

(già precedentemente considerata) ha come vettore gradiente

$$\operatorname{grad} f(\underline{x}) = \begin{cases} \left(\dfrac{1}{x_2}, -\dfrac{x_1}{x_2^2}\right) & \text{se } x_2\neq 0 \\ (0,0) & \text{se } x_1=x_2=0 \end{cases}$$

mentre se $\underline{x}=(x_1,0)$ con $x_1\neq 0$, grad $f(\underline{x})$ non esiste.

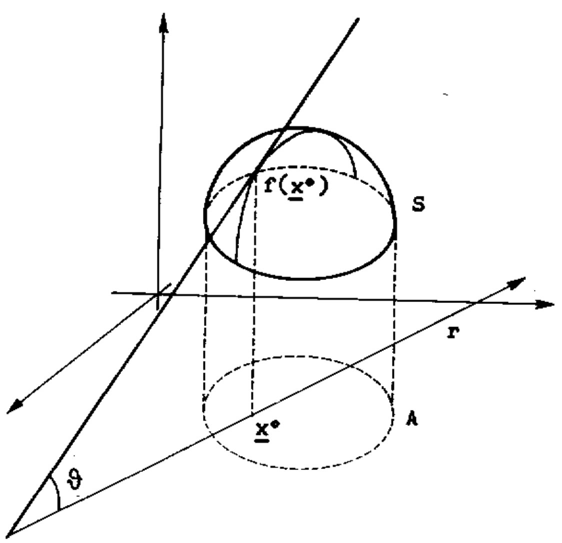
Le regole di calcolo delle derivate prime per funzioni di una variabile sono immediatamente estendibili alle derivate secondo un versore assegnato di funzioni di più variabili. In particolare, se indichiamo con $\underline{k}_1,\dots,\underline{k}_m$ i versori degli assi cartesiani dello spazio R^m, risulta

$$(1.3)\qquad D_{\underline{v}}\underline{f}(\underline{x}) = \sum_{j=1}^{m} D_{\underline{v}} f_j(\underline{x})\underline{k}_j$$

e quindi anche

$$(1.4)\qquad \frac{\partial \underline{f}}{\partial x_i}(\underline{x}) = \sum_{j=1}^{m} \frac{\partial f_j}{\partial x_i}(\underline{x})\underline{k}_j \quad .$$

Se $m=1$, la derivata di f secondo un versore $\underline{v}$ ha il seguente significato geometrico, analogo a quello della derivata prima di una funzione di una sola variabile. Consideriamo, in $R^n \times R$, l'ipersuperficie S diagramma della funzione $y=f(\underline{x})$ e l'iperpiano verticale che ha come traccia in R^n la retta r passante per $\underline{x}^\circ$ e avente la direzione del versore $\underline{v}$. Tale iperpiano taglia S lungo una curva γ. Sia ϑ l'angolo formato dalla retta r e dalla retta tangente a γ nel punto $(\underline{x}^\circ, f(\underline{x}^\circ))$. Vale evidentemente la relazione



$$D_{\underline{v}} f(\underline{x}^\circ) = \operatorname{tang}\vartheta \quad .$$

Facciamo ora alcune importanti osservazioni.

1) Se esiste $D_{\underline{v}}\underline{f}(\underline{x}^\circ)$, la restrizione di $\underline{f}$ alla retta r (uscente da $\underline{x}^\circ$ e parallela al versore $\underline{v}$) è necessariamente continua in $\underline{x}^\circ$.

2) Nella definizione di $D_{\underline{v}}\underline{f}(\underline{x}^\circ)$ intervengono esclusivamente i valori assunti da $\underline{f}$ nei punti di r

in un intorno di $\underline{x}^\circ$; quindi l'esistenza di tale derivata non implica nè la continuità di $\underline{f}$ in $\underline{x}^\circ$ nè l'esistenza delle derivate secondo versori diversi da $\pm\underline{v}$.

3) L'esistenza in un punto $\underline{x}^\circ$ delle derivate secondo tutti i possibili versori di R^n non implica in generale la continuità di $\underline{f}$ in $\underline{x}^\circ$.

Si consideri infatti il seguente controesempio. La funzione $f: R^2 \to R$ così definita:

$$f(x_1,x_2) = \begin{cases} \dfrac{x_1^2}{x_2} & \text{se } x_2 \neq 0 \\ 0 & \text{se } x_2 = 0 \end{cases}$$

ammette in $\underline{x}^\circ=(0,0)$ derivata secondo qualsiasi versore $\underline{v}=(v_1,v_2)$. Precisamente risulta

$$D_{\underline{v}} f(\underline{0}) = \begin{cases} \dfrac{v_1^2}{v_2} & \text{se } v_2 \neq 0 \\ 0 & \text{se } v_2 = 0 \end{cases}$$

eppure f è discontinua nell'origine.

4) L'esistenza in un punto $\underline{x}^\circ$ delle derivate secondo tutti i possibili versori di R^n non implica necessariamente l'esistenza di un legame fra di esse che permetta, conoscendone alcune, di determinare tutte le altre.

Infatti, si considerino la funzione f dell'esempio precedente e la funzione $g: R^2 \to R$ così definita:

$$g(x_1,x_2) = \begin{cases} \dfrac{x_1^2}{x_2} & \text{se } x_2 \neq 0 \\ 2x_1 & \text{se } x_2 = 0 \end{cases} \quad .$$

Risulta

$$D_{\underline{v}} f(\underline{0}) = D_{\underline{v}} g(\underline{0}) \quad \forall\, \underline{v} = (v_1,v_2),\ v_2 \neq 0$$

mentre per $v_2 = 0$ si ha:

$$D_{\underline{v}} f(\underline{0}) = 0 \quad , \quad D_{\underline{v}} g(\underline{0}) = 2 \quad .$$

2. Funzioni differenziabili.

La nozione di derivata secondo un versore è manifestamente incapace di adempiere per le funzioni $\underline{f}: R^n \to R^m$ $(n>1)$ un ufficio analogo a quello della derivata di una funzione $\underline{f}: R \to R^m$: infatti, abbiamo già osservato che, se $n>1$, una funzione può essere derivabile in un punto in qualunque direzione pur essendo discontinua in quel punto.
Per lo studio delle proprietà locali di una funzione $\underline{f}$ di più variabili occorre quindi introdurre un ente, diverso dalle derivate parziali, che sia in grado di farci conoscere l'andamento di $\underline{f}$ in un intorno del punto che ci interessa.
Per avere un orientamento sul modo di procedere, conviene riesaminare ciò che avviene nel caso n=1.
Sia $\underline{f}: R \to R^m$ e supponiamo che esista in un punto $x^\circ \in R$ la derivata prima (finita!) $\underline{f}'(x^\circ)$. Essa ci è utile per studiare il comportamento di $\underline{f}$ in un intorno di x° perchè, in questo caso, $\underline{f}$ è <u>differenziabile</u> in x°, e cioé risulta

$$(2.1) \qquad \Delta \underline{f} = \underline{f}(x^\circ+\Delta x) - \underline{f}(x^\circ) = \underline{f}'(x^\circ)\Delta x + \underline{\varepsilon}(\Delta x)$$

con $\dfrac{\|\underline{\varepsilon}(\Delta x)\|}{\Delta x} \to 0$ per $\Delta x \to 0$. In altri termini, se in x° esiste finita la derivata prima $\underline{f}'(x^\circ)$, è possibile sostituire ad $\underline{f}$, in un intorno abbastanza piccolo di x°, la semplicissima funzione seguente:

$$\underline{g}(x) = \underline{f}(x^\circ) + \underline{f}'(x^\circ)(x-x^\circ) \quad .$$

Ora, è proprio questa proprietà caratteristica delle funzioni di una variabile reale derivabili in un punto con derivata finita che noi dobbiamo estendere al caso di funzioni di più variabili.
Riprendiamo allora in esame la (2.1) e, precisamente, consideriamo il termine $\underline{f}'(x^\circ)\Delta x$. Esso, al variare di Δx in R, definisce una funzione a valori in R^m, lineare (°) in Δx, che dipende (ovviamente!) dalla funzione $\underline{f}$ e dal punto x° considerati. La definizione di funzione differenziabile (quella, ben

---

(°) Cioè additiva e omogenea.

nota, per le funzioni $\underline{f}: R \to R^m$!) può quindi essere data nella forma seguente (equivalente a quella abituale).
Si dice che $\underline{f}: R \to R^m$ è differenziabile in un punto $x^\circ \in R$ se esiste una funzione lineare $\underline{g}: R \to R^m$ tale che risulti

$$\underline{f}(x^\circ+\Delta x) - \underline{f}(x^\circ) = \underline{g}(\Delta x) + \underline{\varepsilon}(\Delta x)$$

con $\dfrac{\|\underline{\varepsilon}(\Delta x)\|}{\Delta x} \to 0$ per $\Delta x \to 0$.
Esaminando questa definizione, è facile vedere che essa ha senso anche per funzioni $\underline{f}: R^n \to R^m$ con $n > 1$ (°).
Introduciamo quindi la seguente

<u>Definizione</u> 2.1 - Sia $\underline{f}: R^n \to R^m$ $(n,m \geq 1)$. Diremo che $\underline{f}$ è differenziabile nel punto $\underline{x}^\circ$ se è possibile determinare un'applicazione lineare $\underline{g}: R^n \to R^m$ tale che risulti

$$(2.2) \qquad \underline{f}(\underline{x}^\circ+\Delta \underline{x}) - \underline{f}(\underline{x}^\circ) = \underline{g}(\Delta \underline{x}) + \underline{\varepsilon}(\Delta \underline{x})$$

con $\dfrac{\|\underline{\varepsilon}(\Delta \underline{x})\|}{\|\Delta \underline{x}\|} \to 0$ per $\Delta \underline{x} \to \underline{0}$.

Ricordiamo ora che una qualsiasi trasformazione lineare $\underline{g}: R^n \to R^m$ $(n,m \geq 1)$ può essere scritta nella forma

$$(2.3) \qquad \underline{y} = L\, \underline{x}$$

dove L è una (m,n) matrice a elementi costanti (indipendenti da $\underline{x}$) e $\underline{x}$ è il generico elemento di R^n scritto come vettore-colonna e che, viceversa, ogni trasformazione di R^n in R^m della forma (2.3) è lineare.
Di conseguenza, alla definizione 2.1 possiamo sostituire la seguente, equivalente

<u>Definizione</u> 2.2 - Sia $\underline{f}: R^n \to R^m$ $(n,m \geq 1)$. Diremo che

---

(°) Anzi, ancor più in generale, per applicazioni $\underline{f}: E \to F$, dove E e F sono due spazi vettoriali normati qualsiasi. In tal caso è però necessario aggiungere l'ulteriore ipotesi che $\underline{g}$ sia continua. Questa ipotesi se $E = R^n$ e $F = R^m$ può essere omessa perchè qualunque applicazione lineare di R^n in R^m è necessariamente continua.

$\underline{f}$ è differenziabile nel punto $\underline{x}^\circ$ se esiste una (m,n)-matrice $L=L(\underline{x})$ i cui elementi dipendono soltanto dal punto $\underline{x}_o$ considerato (e dalla funzione $\underline{f}$) tale che risulti

(2.4) $$\underline{f}(\underline{x}^\circ+\Delta\underline{x})-\underline{f}(\underline{x}^\circ)=L(\underline{x}^\circ)\Delta\underline{x}+\underline{\varepsilon}(\Delta\underline{x})$$

con $\dfrac{\|\underline{\varepsilon}(\Delta\underline{x})\|}{\|\Delta\underline{x}\|}\to 0$ per $\Delta\underline{x}\to\underline{0}$.

Vale il seguente

<u>Teorema</u> 2.1 - Se $\underline{f}$ è differenziabile in $\underline{x}^\circ$, la matrice $L(\underline{x}^\circ)$ è univocamente determinata.
Se infatti esistesse un'altra matrice $\bar{L}(\underline{x}^\circ) = [\bar{\ell}_{j,i}]_{\substack{j=1,\dots,m\\ i=1,\dots,n}} \neq L(\underline{x}^\circ) = [\ell_{j,i}]_{\substack{j=1,\dots,m\\ i=1,\dots,n}}$ tale che

$$f(\underline{x}^\circ+\Delta\underline{x})-f(\underline{x}^\circ)=\bar{L}(\underline{x}^\circ)\Delta\underline{x}+\bar{\varepsilon}(\Delta\underline{x})$$

con $\dfrac{\|\bar{\varepsilon}(\Delta\underline{x})\|}{\|\Delta\underline{x}\|}\to 0$ per $\Delta\underline{x}\to\underline{0}$, si avrebbe

$$\bar{L}(\underline{x}^\circ)\Delta\underline{x}-L(\underline{x}^\circ)\Delta\underline{x}=\underline{\varepsilon}(\Delta\underline{x})-\bar{\underline{\varepsilon}}(\Delta\underline{x})=o(\|\Delta\underline{x}\|)$$

per $\Delta\underline{x}\to\underline{0}$. Ma allora, assumendo $\Delta\underline{x}=\Delta x_i\underline{e}_i$ $(i=1,\dots,n)$ si otterrebbe

$$(\bar{\ell}_{j,i}-\ell_{j,i})\Delta x_i = o(\Delta x_i) \qquad (j=1,\dots,m)$$

per $\Delta x_i\to 0$, assurdo se $\bar{\ell}_{j,i}\neq\ell_{j,i}$.

si suppone di aver fissato una base in $\mathbb{R}^2$

Il primo termine del secondo membro di (2.4) viene chiamato <u>differenziale</u> (primo) di $\underline{f}$ in $\underline{x}^\circ$ e indicato con $d\underline{f}(\underline{x}^\circ)$. Se $\underline{f}$ è l'identità, $d\underline{f}=\Delta\underline{x}$. Possiamo quindi scrivere :

(2.5) $$d\underline{f}(\underline{x}^\circ) = L(\underline{x}^\circ)d\underline{x} \quad .$$

Se $m=1$, l'iperpiano di $R^n\times R$ di equazione

$$y - f(\underline{x}^\circ) = L(\underline{x}^\circ)(\underline{x}-\underline{x}^\circ)$$

viene chiamato <u>iperpiano tangente</u> alla superficie diagramma della funzione f nel punto $\underline{x}^\circ$.

Consideriamo, ad esempio, la funzione $f:R^2\to R$ definita da $f(x_1,x_2) = x_1x_2$. Risulta :

$$\Delta f = f(\underline{x}^\circ+\Delta\underline{x})-f(\underline{x}^\circ) = x_2^\circ\Delta x_1 + x_1^\circ\Delta x_2+\Delta x_1\Delta x_2 .$$

Poichè

$$\frac{|\Delta x_1\Delta x_2|}{\|\Delta\underline{x}\|} = \frac{|\Delta x_1\Delta x_2|}{\sqrt{\Delta x_1^2+\Delta x_2^2}} \leq |\Delta x_2| \to 0 \quad \text{per} \quad \Delta\underline{x}\to\underline{0} ,$$

posto $L(\underline{x}^\circ) = [x_2^\circ, x_1^\circ]$ vale la (2.4) e f è differenziabile in ogni punto di R^2 .
Risulta inoltre

$$df(\underline{x}^\circ) = x_2^\circ dx_1 + x_1^\circ dx_2 \qquad \forall\,\underline{x}^\circ\in R^2 .$$

Il piano tangente al diagramma di f nel generico punto $\underline{x}^\circ$ ha equazione

$$y - x_1^\circ x_2^\circ = L(\underline{x}^\circ)\,(\underline{x}-\underline{x}^\circ)$$

e cioè

$$y-x_1^\circ x_2^\circ = x_2^\circ(x_1-x_1^\circ)+x_1^\circ(x_2-x_2^\circ) .$$

Osserviamo infine che $\underline{f}:R^n\to R^m$ è differenziabile in $\underline{x}^\circ$ se e solo se ogni sua componente è differenziabile in $\underline{x}^\circ$. Inoltre, è facile dimostrare che se f e g sono funzioni reali differenziabili in $\underline{x}^\circ$, anche f+g e f·g sono differenziabili in $\underline{x}^\circ$ e lo stesso dicasi per f/g purché sia $g(\underline{x}^\circ)\neq 0$.

3. <u>Prime proprietà delle funzioni differenziabili</u>.

Converrà premettere ai teoremi di questo paragrafo un Lemma che ci sarà utile qui e nel seguito.

(*) <u>Lemma</u> 3.1 - Sia $L=[\ell_{j,i}]_{\substack{j=1,\dots,m\\ i=1,\dots,n}}$ una (m,n)-matrice a elementi reali e si consideri l'applicazione lineare di R^n in R^m $\underline{y}=L\underline{x}$ definita da L . Allora, per ogni $\underline{x}$ di R^n, vale la disuguaglianza

$$\|\underline{y}\| \leq m\, n\ell\|\underline{x}\|$$

dove $\ell = \sup\limits_{\substack{j=1,\dots,m\\ i=1,\dots,n}} |\ell_{j,i}|$.

Infatti

$$\|\underline{y}\|=\|L\underline{x}\|=\left\{\sum_{j=1}^{m}\left|\sum_{i=1}^{n}\ell_{j,i}\Delta x_i\right|^2\right\}^{\frac12} \leq$$

$$\leq\left\{\sum_{j=1}^{m}\left\{\sum_{i=1}^{n}\ell\|\Delta\underline{x}\|\right\}^2\right\}^{\frac12} \leq mn\ell\|\Delta\underline{x}\| \quad .$$

(*) *Vedere le varie norme di matrici.*

Teorema 3.1 - Se $\underline{f}: R^n \to R^m$ $(n,m \geq 1)$ è differenziabile in $\underline{x}^\circ$, essa è continua in $\underline{x}^\circ$.
Infatti, da (2.4) e dal Lemma 3.1, posto $L(\underline{x}^\circ) = [\ell_{j,i}]$ si ricava :

$$(3.1) \quad \| \underline{f}(\underline{x}^\circ + \Delta \underline{x}) - \underline{f}(\underline{x}^\circ) \| \leq \| L(\underline{x}^\circ) \Delta \underline{x} \| + \| \underline{\varepsilon}(\Delta \underline{x}) \|$$

$$\leq m\, n\, \ell \, \| \Delta \underline{x} \| + \| \varepsilon(\Delta \underline{x}) \| = O(\| \Delta \underline{x} \|) \quad \text{per} \quad \Delta \underline{x} \to \underline{0}.$$

Teorema 3.2 - Se $\underline{f}: R^n \to R^m$ è differenziabile in $\underline{x}^\circ$ esistono in $\underline{x}^\circ$ le derivate di $\underline{f}$ secondo qualsiasi versore $\underline{v}$ di R^n, inoltre risulta

$$(3.2) \qquad D_{\underline{v}} \underline{f}(\underline{x}^\circ) = L(\underline{x}^\circ) \underline{v} \quad .$$

Infatti, da (2.4) si ottiene

$$(3.3) \qquad \frac{\underline{f}(\underline{x}^\circ + t\underline{v}) - \underline{f}(\underline{x}^\circ)}{t} = L(\underline{x}^\circ)\underline{v} + \frac{\underline{\varepsilon}(t\underline{v})}{t} \quad .$$

Poichè $\dfrac{\| \underline{\varepsilon}(t\underline{v}) \|}{|t|} = \dfrac{\| \varepsilon(t\underline{v}) \|}{\| t\underline{v} \|} \to 0$ per $t \to 0$,

da (3.3) segue che esiste

$$D_{\underline{v}} \underline{f}(\underline{x}^\circ) = \lim_{t \to 0} \frac{\underline{f}(\underline{x}^\circ + t\underline{v}) - \underline{f}(\underline{x}^\circ)}{t} = L(\underline{x}^\circ)\underline{v}$$

e cioè l'asserto.

Dalla formula (3.2) si ricava, fra l'altro, il seguente risultato.

Teorema 3.3 - Sia $\underline{f}$ differenziabile in $\underline{x}^\circ$. Allora la matrice $L(\underline{x}^\circ)$ coincide con la matrice jacobiana $J(\underline{f}, \underline{x}^\circ)$ di $\underline{f}$ nel punto $\underline{x}^\circ$.
Infatti, sia $L(\underline{x}^\circ) = [\ell_{j,i}]_{\substack{j=1,\dots,m \\ i=1,\dots,n}}$. Posto $\underline{v} = \underline{e}_i$, da (3.2) si ricava

$$\frac{\partial \underline{f}}{\partial x_i}(\underline{x}^\circ) = D_{\underline{e}_i} \underline{f}(\underline{x}^\circ) = L(\underline{x}^\circ)\underline{e}_i = [\ell_{j,i}]_{j=1,\dots,m} \quad .$$

e cioè
$$\ell_{j,i} = \frac{\partial f_j}{\partial x_i}(\underline{x}^\circ) \qquad j=1,\dots,m; \; i=1,\dots,n$$

c.d.d.

Dalla (2.4) e dal teorema 3.3 si ricava il

Corollario 3.1 - $\underline{f}: R^n \to R^m$ $(n,m \geq 1)$ è differenziabile in $\underline{x}^\circ$ se e soltanto se esistono tutte le derivate parziali prime $\partial f_j / \partial x_i$ $(\underline{x}^\circ)$ $(j=1,\dots,m;\ i=1,\dots,n)$ e se inoltre

$$\underline{f}(\underline{x}^\circ + \Delta \underline{x}) - \underline{f}(\underline{x}^\circ) - J(\underline{f}, \underline{x}^\circ) \Delta \underline{x}$$

è infinitesimo di ordine superiore al primo rispetto a $\| \Delta \underline{x} \|$.

Il Corollario 3.1 viene sovente applicato per stabilire se una funzione è differenziabile oppure no in un punto.

E s e m p i

1) La funzione $f: R^2 \to R$ definita da $f(x_1, x_2) = \sqrt{|x_1 x_2|}$ non è differenziabile nei punti degli assi coordinati.

Infatti, se $\underline{x}^\circ = (0, x_2)$ con $x_2 \neq 0$, la derivata parziale $\dfrac{\partial f}{\partial x_1}(\underline{x}^\circ)$ non esiste e quindi f non può essere differenziabile in tale punto.
Per motivi di simmetria f non può essere differenziàbile nei punti della forma $(x_1, 0)$ con $x_1 \neq 0$.
Se $\underline{x}^\circ = \underline{0}$ si ha $\dfrac{\partial f}{\partial x_1}(\underline{0}) = \dfrac{\partial f}{\partial x_2}(\underline{0}) = 0$ e $f(\underline{0} + \Delta \underline{x}) - f(\underline{0}) - J(f; \underline{0}) \Delta \underline{x} = \sqrt{|\Delta x_1 \Delta x_2|}$.
Quindi nel punto $\underline{x}^\circ = \underline{0}$ f non è differenziabile, perchè

$$\frac{\sqrt{|\Delta x_1 \Delta x_2|}}{\| \Delta x \|} = \frac{\sqrt{|\Delta x_1 \Delta x_2|}}{\sqrt{\Delta x_1^2 + \Delta x_2^2}}$$

non tende a zero per $\Delta \underline{x} \to \underline{0}$. (Basta infatti assumere $\Delta x_1 = \Delta x_2 \to 0$ perchè il rapporto precedente tenda a $\frac{1}{\sqrt{2}}$).

2) La funzione $f: R^2 \to R$ definita da $f(x_1, x_2) = (x_1 + 1)\sqrt{|x_2|}$ é differenziabile nel punto $\underline{x}^\circ = (-1, 0)$.

Infatti in tale punto entrambe le derivate parziali prime sono nulle.
Quindi $f(\underline{x}^\circ + \Delta \underline{x}) - f(\underline{x}^\circ) - J(f; \underline{x}^\circ) \Delta \underline{x} = \Delta x_1 \sqrt{|\Delta x_2|}$.
Consideriamo il rapporto $\dfrac{\Delta x_1 \sqrt{|\Delta x_2|}}{\| \Delta \underline{x} \|}$.

Poichè risulta
$$\frac{|\Delta x_1|\sqrt{|\Delta x_2|}}{\|\Delta \underline{x}\|} = \frac{|\Delta x_1|\sqrt{|\Delta x_2|}}{\sqrt{\Delta x_1^2+\Delta x_2^2}} < \sqrt{|\Delta x_2|} \to 0$$

per $\Delta \underline{x} \to \underline{0}$, f è differenziabile in (-1,0).

Dai teoremi 3.2 e 3.3 si ricavano anche i seguenti corollari.

<u>Corollario</u> 3.2 - Se $\underline{f}: R^n \to R^m$ $(n,m \geq 1)$ è differenziabile in $\underline{x}^\circ$, risulta
$$d\underline{f} = J(\underline{f};\underline{x}^\circ)d\underline{x}$$
e, in particolare, se $m=1$
$$df = \text{grad } f(\underline{x}^\circ)d\underline{x} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\underline{x}^\circ)dx_i .$$

<u>Corollario</u> 3.3 - Se $\underline{f}: R^n \to R^m$ è differenziabile in $\underline{x}^\circ$, per ogni versore $\underline{v}$ di R^n risulta
$$(3.4) \qquad D_{\underline{v}}\underline{f}(\underline{x}^\circ) = J(\underline{f};\underline{x}^\circ)\underline{v} \; ,$$
e in particolare, se $m=1$:
$$D_{\underline{v}}f(\underline{x}^\circ) = \text{grad } f(\underline{x}^\circ)\underline{v} \; .$$

Ad esempio, per la funzione $\underline{f}: R^2 \to R^2$ definita da $\underline{f}(x_1,x_2)=(x_1\text{sen } x_2, x_1^2)$ risulta
$$D_{\underline{v}}\underline{f}(\underline{x}) = \begin{bmatrix} \text{sen } x_2 & x_1\cos x_2 \\ 2x_1 & 0 \end{bmatrix} \cdot \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = (v_1 \text{sen } x_2 + v_2 x_1 \cos x_2, 2v_1x_1)$$
e, in particolare, per la funzione $f_1: R^2 \to R$ definita da $f_1(x_1,x_2) = x_1 \text{sen } x_2$ si ha
$$D_{\underline{v}}f_1(\underline{x}) = v_1 \text{sen } x_2 + v_2 x_1 \cos x_2 \; .$$

<u>Osservazioni</u> - 1) La (3.4) si può anche scrivere nella forma
$$(3.5) \qquad D_{\underline{v}}\underline{f}(\underline{x}^\circ) = \sum_{i=1}^{n} \frac{\partial \underline{f}}{\partial x_i}(\underline{x}^\circ)v_i \; .$$

2) la continuità di $\underline{f}$ in $\underline{x}^\circ$ e l'esistenza per tutti i versori $\underline{v}$ di R^n delle derivate $D_{\underline{v}}\underline{f}(\underline{x}^\circ)$ soddisfacenti la (3.4) non bastano ad assicurare la differenziabilità di $\underline{f}$ in $\underline{x}^\circ$.

Si consideri infatti il seguente controesempio. Sia
$$f(x_1,x_2) = \begin{cases} \sqrt{x_1^2+x_2^2}\left(1-e^{-\frac{x_1^2+x_2^2}{|x_1|}}\right) & \text{se } x_1 \neq 0 \\ 0 & \text{se } x_1 = 0 \end{cases} .$$

Essa è continua nell'origine; infatti:
$$|f(x_1,x_2)| \leq \sqrt{x_1^2+x_2^2} \to 0 \quad \text{per } (x_1,x_2) \to (0,0) \; .$$

Inoltre per ogni versore $\underline{v}$ di R^2 si ha
$$\frac{f(\underline{o}+t\underline{v})-f(\underline{o})}{t} = \begin{cases} \left(1-e^{-\frac{t^2}{|tv_1|}}\right)\text{sgn } t & \text{se } v_1 \neq 0 \\ 0 & \text{se } v_1 = 0 \end{cases}$$

quindi $D_{\underline{v}}f(\underline{O})=0$ e vale (3.4). Eppure f non è differenziabile nell'origine; perchè il rapporto
$$\frac{f(\underline{o}+\Delta\underline{x})-f(\underline{o})-J(f,\underline{o})\Delta\underline{x}}{\|\Delta\underline{x}\|} = \begin{cases} 1-e^{-\frac{\|\Delta\underline{x}\|^2}{|\Delta x_1|}} & \text{se } \Delta x_1 \neq 0 \\ 0 & \text{se } \Delta x_1 = 0 \end{cases}$$
non tende a zero. Infatti, basta assumere $\Delta x_1 = \Delta x_2^2 \to 0$ per ottenere
$$1-e^{-\frac{\|\Delta\underline{x}\|^2}{|\Delta x_1|}} \to -e^{-1} .$$

4. <u>Composizione di funzioni differenziabili</u>.

<u>Teorema</u> 4.1 - Sia $\underline{f}: R^n \to R^m$ una funzione differenziabile in $\underline{x}^\circ$; $\underline{g}: R^m \to R^p$ una funzione differenziabile nel punto $\underline{y}^\circ = \underline{f}(\underline{x}^\circ)$ $(m,n,p \geq 1)$. Allora la funzione composta $\underline{g} \circ \underline{f}: R^n \to R^p$ è differenziabile in $\underline{x}^\circ$ e

(4.1) $$J(\underline{g}o\underline{f};\underline{x}^\circ)=J(\underline{g};\underline{y}^\circ)\cdot J(\underline{f};\underline{x}^\circ) \quad .$$

Infatti, poniamo, per semplicità di scrittura, $J(\underline{f};\underline{x}^\circ)=A$, $J(\underline{g};\underline{y}^\circ)=B$. Si ha

(4.2) $$\underline{f}(\underline{x}^\circ+\Delta\underline{x})-\underline{f}(\underline{x}^\circ)=A\Delta\underline{x}+o(\|\Delta\underline{x}\|)$$

(4.3) $$\underline{g}(\underline{y}^\circ+\Delta\underline{y})-\underline{g}(\underline{y}^\circ)=B\Delta\underline{y}+o(\|\Delta\underline{y}\|) \quad .$$

Assumiamo $\Delta\underline{y}=\underline{f}(\underline{x}^\circ+\Delta\underline{x})-\underline{f}(\underline{x}^\circ)$. Da (4.2) si ricava, ricordando il Lemma 3.1

(4.4) $$\Delta\underline{y} = A\Delta\underline{x} + o(\|\Delta\underline{x}\|) = O(\|\Delta\underline{x}\|) \quad .$$

Di qui, sostituendo in 4.3, si ottiene

$$\underline{g}o\underline{f}(\underline{x}^\circ+\Delta\underline{x})-\underline{g}o\underline{f}(\underline{x}^\circ)=BA\Delta\underline{x}+B\, o(\|\Delta\underline{x}\|)+o(O(\|\Delta\underline{x}\|))$$

da cui, ricordando ancora il Lemma 3.1, si ha

$$\underline{g}o\underline{f}(\underline{x}^\circ+\Delta\underline{x})-\underline{g}o\underline{f}(\underline{x}^\circ)=BA\Delta\underline{x}+o(\|\Delta\underline{x}\|)$$

e cioè l'asserto.

<u>Osservazioni</u> - 1) Da (4.1) si ricava

(4.5) $$d\underline{g}o\underline{f}(\underline{x}^\circ)=J(\underline{g};\underline{y}^\circ)J(\underline{f};\underline{x}^\circ)d\underline{x}$$

e quindi, posto per semplicità di scrittura $\underline{z}=\underline{g}o\underline{f}$ si ha

(4.6) $$\frac{\partial z_j}{\partial x_i}(\underline{x}^\circ)=\sum_{k=1}^{m}\frac{\partial g_j}{\partial y_k}(\underline{f}(\underline{x}^\circ))\cdot\frac{\partial f_k}{\partial x_i}(\underline{x}^\circ) \quad .$$

In particolare, se $p=1$ si ottiene

$$\frac{\partial z}{\partial x_i}(\underline{x}^\circ)=\sum_{k=1}^{m}\frac{\partial g}{\partial y_k}(\underline{f}(\underline{x}^\circ))\cdot\frac{\partial f_k}{\partial x_i}(\underline{x}^\circ) \quad .$$

Più in particolare ancora, se $n=p=1$ si ricava

(4.7) $$\frac{dz}{dx}(x^\circ)=\text{grad }g(\underline{f}(x^\circ))\cdot\frac{d\underline{f}}{dx}(x^\circ) = \sum_{k=1}^{m}\frac{\partial g}{\partial y_k}(\underline{f}(x^\circ))\cdot\frac{df_k}{dx}(x^\circ) \quad .$$

Infine, se $n=m=p=1$ si ha

$$\frac{dz}{dx}(x^\circ) = \frac{dg}{dy}(f(x^\circ))\cdot\frac{df}{dx}(x^\circ) \quad ,$$

e cioè la regola di derivazione delle funzioni composte di una sola variabile reale.

Calcoliamo, ad esempio, le derivate parziali prime nell'origine della funzione $\underline{g}o\underline{f}$ composta delle due funzioni

$$\underline{f}(x_1,x_2)=(x_1-1\ ,x_1+2x_2)$$

$$\underline{g}(y_1,y_2)=((y_1+1)\sqrt{|y_2|},\ y_1^2+y_2) \quad .$$

Le due componenti di $\underline{f}$ sono differenziabili nell'origine (perchè somma di funzioni differenziabili). Essendo la funzione $\underline{g}$ differenziabile nel punto $\underline{y}^\circ=\underline{f}(\underline{0})=(-1,0)$ (vedasi l'esempio 2 in calce al Corollario 3.1), anche $\underline{z}=\underline{g}o\underline{f}$ è differenziabile nell'origine. Inoltre si ha :

$$\frac{\partial f_1}{\partial x_1}(\underline{0})=\frac{\partial f_2}{\partial x_1}(\underline{0})=1\ ,\quad \frac{\partial f_1}{\partial x_2}(\underline{0})=0\ ,\quad \frac{\partial f_2}{\partial x_2}(\underline{0})=2\ ;$$

$$\frac{\partial g_1}{\partial y_1}(\underline{y}^\circ)=0\ ,\quad \frac{\partial g_2}{\partial y_1}(\underline{y}^\circ)=-2\ ,\quad \frac{\partial g_1}{\partial y_2}(\underline{y}^\circ)=0\ ,\quad \frac{\partial g_2}{\partial y_2}(\underline{y}^\circ)=1$$

e quindi da (4.6) si ottiene

$$\frac{\partial z_1}{\partial x_1}(\underline{0})=\frac{\partial z_1}{\partial x_2}(\underline{0})=0\ ,\quad \frac{\partial z_2}{\partial x_1}(\underline{0})=-1\ ,\quad \frac{\partial z_2}{\partial x_2}(\underline{0})=2 \quad .$$

2) Siano A_1 e A_2 due aperti, rispettivamente di R^n e R^m $(n,m\geq 1)$ e siano $\underline{f}:A_1\to A_2$ differenziabile in ogni punto di A , $\underline{g}:A_2\to R^p$ $(p\geq 1)$ differenziabile in ogni punto di A_2 della forma $\underline{y}=\underline{g}(\underline{x})$. Allora la funzione $\underline{z}=\underline{g}o\underline{f}$ è differenziabile in ogni punto di A_1 e da (4.5) segue

$$d\underline{z} = J(\underline{g};\underline{y})J(\underline{f};\underline{x})d\underline{x}=J(\underline{g};\underline{y})d\underline{y} \quad .$$

Dunque il differenziale dz della funzione $\underline{g}(\underline{y})$ ha l'espressione $d\underline{z}=J(\underline{g};\underline{y})d\underline{y}$ tanto se il vettore $\underline{y}$ è variabile indipendente, quanto se esso è a sua volta funzione di un altro vettore variabile $\underline{x}$. Sussiste cioè anche nel caso di funzioni di più variabili reali il principio della <u>invarianza formale</u> del differnziale di fronte alla scelta delle variabili indipendenti, principio che si è già visto valere nel caso di funzioni di una sola variabile reale.

5. Una condizione sufficiente per la differenziabilità.

La differenziabilità di una funzione $\underline{f}: R^n \to R^m$ $(n, m \geq 1)$ in un aperto A di R^n $(n \geq 1)$ implica l'esistenza delle derivate prime di $\underline{f}$ in A ma non implica in generale la loro continuità.

Si consideri, come contro-esempio, la funzione $f: R \to R$ definita da

$$f(0)=0 \quad , \quad f(x)=x^2 \operatorname{sen} \frac{1}{x^2} \quad \text{se } x \neq 0 \quad .$$

Essa è differenziabile in ogni punto $x \in R$ e risulta

$$f'(0)=0 \quad , \quad f'(x)=2x \operatorname{sen} \frac{1}{x^2} - \frac{2}{x} \cos \frac{1}{x^2} \quad \text{se } x \neq 0 \quad .$$

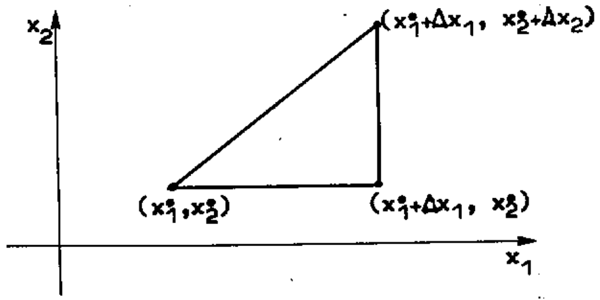
Quindi f' non è continua nell'origine. Si osservi che f' non è neppure limitata in alcun intorno dell'origine.

Invece, la continuità delle derivate prime in A implica necessariamente la differenziabilità di $\underline{f}$. Vale infatti il

Teorema 5.1 - Se in un intorno del punto $\underline{x}^\circ$ esistono tutte le derivate parziali prime di $\underline{f}$ e sono continue in $\underline{x}^\circ$, allora $\underline{f}$ è differenziabile in $\underline{x}^\circ$.

Basta evidentemente dimostrare il teorema per m=1. Possiamo scrivere

$$(5.1) \qquad f(\underline{x}^\circ+\Delta\underline{x})-f(\underline{x}^\circ)=\sum_{i=1}^{n}\left(g_i(x_i^\circ+\Delta x_i)-g_i(x_i^\circ)\right)$$



dove si è posto, per semplicità di scrittura

$$g_i(t)=f(x_1^\circ+\Delta x_1, \ldots, x_{i-1}^\circ+\Delta x_{i-1}, t, x_{i+1}^\circ, \ldots, x_n^\circ) \ .$$

La funzione ausiliaria g_i $(i=1,\ldots,n)$ è differenziabile nel punto x_i° per $\|\Delta\underline{x}\|$ abbastanza piccola e risulta

$$\frac{dg_i}{dt}(x_i^\circ)=\frac{\partial f}{\partial x_i}(x_1^\circ+\Delta x_1,\ldots,x_{i-1}^\circ+\Delta x_{i-1},x_i^\circ,x_{i-1}^\circ,\ldots,x_n^\circ)$$

e quindi, per la continuità di $\frac{\partial f}{\partial x_i}$ in $\underline{x}^\circ$

$$(5.2) \qquad \frac{dg_i}{dt}(x_i^\circ)=\frac{\partial f}{\partial x_i}(\underline{x}^\circ)+o(1) \qquad \text{per } \Delta\underline{x}\to\underline{0} \ .$$

D'altra parte, è anche, per $\Delta x_i \to 0$

$$g_i(x_i^\circ+\Delta x_i)-g_i(x_i^\circ)=\frac{dg}{dt}(x_i^\circ)\Delta x_i+o(\Delta x_i)$$

da cui, per la (5.2), si ricava

$$g_i(x_i^\circ+\Delta x_i)-g_i(x_i^\circ)=\frac{\partial f}{\partial x_i}(\underline{x}^\circ)\Delta x_i+o(\|\Delta\underline{x}\|) \qquad \text{per } \Delta\underline{x}\to\underline{0}$$

Sostituendo in (5.1) si ottiene infine

$$f(\underline{x}^\circ+\Delta\underline{x})-f(\underline{x}^\circ)=\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(\underline{x}^\circ)\,\Delta x_i+o(\|\Delta\underline{x}\|) = \\ =\operatorname{grad} f(\underline{x}^\circ)\Delta\underline{x}+o(\|\Delta\underline{x}\|)$$

e cioè l'asserto.

Corollario 5.1 - Se in un aperto A esistono e sono continue tutte le derivate parziali prime di $\underline{f}$, allora $\underline{f}$ è differenziabile in A. Inoltre esistono e sono continue in A le derivate di $\underline{f}$ in qualsiasi direzione.
Basta infatti ricordare che $D_{\underline{v}}\underline{f}(\underline{x})=J(\underline{f};\underline{x})\underline{v}$ e che gli elementi della matrice jacobiana sono continui in A .

Ad esempio, la funzione $f(x_1,x_2)=\frac{x_1}{x_2^2+1}$ è differenziabile in R^2 perché in ogni punto di R^2 esistono

$$\frac{\partial f}{\partial x_1}=\frac{1}{x_2^2+1} \qquad , \qquad \frac{\partial f}{\partial x_2}=\frac{-2x_1x_2}{(x_2^2+1)^2}$$

e sono continue in R^2 . Inoltre per ogni versore $\underline{v}$ di R^2 si ha

$$D_{\underline{v}}f=\frac{v_1}{x_2^2+1}-\frac{2x_1x_2v_2}{(x_2^2+1)^2} \ .$$

Una funzione che soddisfi in un aperto A le ipotesi del corollario 5.1, cioé una funzione dotata di derivate parziali prime continue in A, viene chiamata differenziabile con continuità in A, o di classe $\mathscr{C}^{(1)}(A)$.

6. Teorema dell'incremento finito: il caso m=1.

Vale, senza modifiche, nel caso in cui lo spazio di partenza ha un Banach.

Per le funzioni reali di più variabili reali vale il seguente teorema dell'incremento finito, del tutto analogo a quello valido per funzioni reali di una variabile reale.

Teorema 6.1 - La funzione $f: R^n \to R$ $(n \geq 1)$ sia differenziabile in un aperto A e il segmento chiuso di estremi $\underline{x}^\circ$, $\underline{x}^\circ + \underline{h}$ appartenga ad A. Allora esiste $\vartheta = \vartheta(\underline{x}^\circ, \underline{h})$, $0 < \vartheta < 1$, tale che

$$f(\underline{x}^\circ+\underline{h}) - f(\underline{x}^\circ) = \text{grad } f(\underline{x}^\circ + \vartheta\underline{h}) \cdot \underline{h} \quad . \tag{6.1}$$

Infatti, se $n=1$, la (6.1) diventa

$$f(x^\circ+h) - f(x^\circ) = f'(x^\circ+\vartheta h) \cdot h \tag{6.2}$$

e il teorema 6.1 si riduce, in questo caso, al teorema dell'incremento finito per funzioni $f: R \to R$, che supponiamo noto.
Se $n > 1$, consideriamo la funzione ausiliaria $g: R \to R$ $g(t) = f(\underline{x}^\circ + t\underline{h})$. Essa, essendo composta della funzione $\varphi: R \to R^n$ $\varphi(t) = \underline{x}^\circ + t\underline{h}$ (differenziabile per ogni t) e della funzione f, risulta continua in $[0,1]$, differenziabile in $(0,1)$ e inoltre da (4.7) si ricava

$$g'(t) = \text{grad } f(\underline{x}^\circ + t\underline{h}) \cdot \underline{h} \qquad \forall t \in (0,1).$$

Dal teorema dell'incremento finito nel caso $n=1$ applicato a g si ottiene allora

$$f(\underline{x}^\circ+\underline{h}) - f(\underline{x}^\circ) = g(1) - g(0) = g'(\vartheta) \cdot 1 = \text{grad } f(\underline{x}^\circ + \vartheta\underline{h}) \cdot \underline{h}$$

con ϑ opportuno, $0 < \vartheta < 1$, c.d.d.

Osservazione - Il teorema 6.1 vale anche con le meno restrittive ipotesi seguenti:

a) f sia definita in un insieme contenente il segmento chiuso S di estremi $\underline{x}^\circ$, $\underline{x}^\circ + \underline{h}$;

b) la restrizione di f a S sia continua;

c) f sia differenziabile in ogni punto interno a S.

La dimostrazione del teorema in questa forma più generale è identica alla precedente.

Dal teorema 6.1 si ricava, fra l'altro, la seguente caratterizzazione delle funzioni costanti in un aperto connesso.

Teorema 6.2 - Condizione necessaria e sufficiente affinchè una funzione $\underline{f}: R^n \to R^m$ $(n, m \geq 1)$, differenziabile in un aperto connesso A, sia costante in A è che

$$d\underline{f}(\underline{x}) = \underline{0} \qquad \forall \underline{x} \in A \quad . \tag{6.2}$$

Infatti, se $\underline{f}$ è costante in A, ovviamente $d\underline{f} = \underline{0}$ in A. Viceversa, per dimostrare che, se vale (6.2), $\underline{f}$ è costante in A, basta considerare il caso m=1. Sia $\underline{x}^\circ \in A$ arbitrario e siano, rispettivamente

$$A_1 = \{\underline{x} \in A \ : \ f(\underline{x}) = f(\underline{x}^\circ)\}$$
$$A_2 = \{\underline{x} \in A \ : \ f(\underline{x}) \neq f(\underline{x}^\circ)\} \ .$$

A_1 è aperto. Infatti, sia $\underline{x} \in A_1$ e sia d la sua distanza dalla frontiera di A. Se $\|\underline{h}\| < d$, da (6.1) si ricava $f(\underline{x}+\underline{h}) - f(\underline{x}) = 0$. Quindi, se $\|\underline{h}\| < d$, $\underline{x}+\underline{h} \in A_1$.
A_2 è aperto. Infatti, se $\underline{x} \in A_2$ si ha $|f(\underline{x}) - f(\underline{x}^\circ)| > a > 0$ e, per la continuità di f, esiste un intorno di $\underline{x}$ tale che per ogni suo punto $\underline{y}$ risulta $\|f(\underline{y}) - f(\underline{x}^\circ)\| > a/2$. Quindi tale intorno è contenuto in A_2.
Ma $A = A_1 \cup A_2$ è connesso; perciò uno dei due aperti A_1, A_2 deve essere vuoto. Poichè A_1 contiene $\underline{x}^\circ$ si ha allora $A_2 = \emptyset$, c.d.d.

7. Teorema dell'incremento finito: il caso $m>1$.

Il teorema dell'incremento finito non è estendibile nella forma abituale al caso $m>1$, neanche quando $n=1$.

Si consideri infatti come contro-esempio la funzione $\underline{f}:R\to R^2$, differenziabile per ogni x, definita da $\underline{f}(x)=(\mathrm{sen}\ x, \cos x)$. Se valesse una formula analoga a (6.1) dovrebbe essere

$$\underline{f}(2\pi)-\underline{f}(0)=\underline{f}'(2\pi\vartheta)\cdot 2\pi$$

per un ϑ opportuno, $0<\vartheta<1$. Ma $\underline{f}(2\pi)=\underline{f}(0)$ e quindi dovrebbe essere $\underline{f}'(2\pi\vartheta)=\underline{0}$, assurdo perchè $\underline{f}'(x)=(\cos x, -\mathrm{sen}\ x)\neq\underline{0}\ \forall x\in R$.

E' però valida, anche nel caso $m>1$, la seguente forma attenuata dal teorema.

Questo teorema va sostituito con [H, Teorema 1.1.1 e seguenti]

Teorema 7.1 - La funzione $\underline{f}:R^n\to R^m$ $(n,m\geq 1)$ sia differenziabile in un aperto A. Allora, se il segmento chiuso di estremi $\underline{x}^\circ$, $\underline{x}^\circ+\underline{h}$ appartiene ad A, vale la limitazione

$$\|\underline{f}(\underline{x}^\circ+\underline{h})-\underline{f}(\underline{x}^\circ)\|\leq m\,n\,M\,\|\underline{h}\| \tag{7.1}$$

dove

$$M=\mathrm{Sup}\left|\frac{\partial f_j}{\partial x_i}(\underline{x})\right| \tag{7.2}$$

al variare di j fra 1 e m, di i fra 1 e n e di $\underline{x}$ sul segmento di estremi $\underline{x}^\circ$, $\underline{x}^\circ+\underline{h}$.

Infatti dal teorema 6.1 si ricava

$$\left|f_j(\underline{x}^\circ+\underline{h})-f_j(\underline{x}^\circ)\right|\leq\|\mathrm{grad}\ \underline{f}\|\cdot\|\underline{h}\|\leq n\,M\,\|\underline{h}\|$$

per $j=1,\dots,m$ e quindi

$$\|\underline{f}(\underline{x}^\circ+\underline{h})-\underline{f}(\underline{x}^\circ)\|\leq m\,n\,M\,\|\underline{h}\| \qquad \text{c.d.d.}$$

Osservazioni - 1) Anche le ipotesi del teorema 7.1 possono venire attenuate nello stesso modo di quelle del teorema 6.1.

2) Se n o m sono maggiori di uno, la costante $m\,n\,M$ che compare nel secondo membro di (7.1) non è, in generale, la migliore possibile. Noi ci

limitiamo però ad enunciare, per semplicità, il teorema in questa forma.

Dal teorema 7.1 si ricava come corollario il seguente, che estende al caso di funzioni di più variabili una condizione sufficiente per la lipschitzianità, già dimostrata per funzioni $f:R\to R$ (°).

Cfr. Rudin Es. 7, p. 239

Teorema 7.2 - Sia C un convesso di R^n. La funzione $\underline{f}:C\to R^m$ sia continua in C e differenziabile in ogni punto interno a C. Inoltre esista una costante K tale che

$$\left|\frac{\partial f_j}{\partial x_i}\right|\leq K \qquad (j=1,\dots,m,\quad i=1,\dots,n)$$

per ogni $\underline{x}$ interno a C. Allora $\underline{f}$ è lipschitziana in C con costante di Lipschitz $m\,n\,K$.

Infatti, siano $\underline{x}_1, \underline{x}_2\in C$. Allora dal teorema 7.1 (tenuto conto dell'osservaz. 1) si ricava

$$\|\underline{f}(\underline{x}_2)-\underline{f}(\underline{x}_1)\|\leq m\,n\,M\,\|\underline{x}_2-\underline{x}_1\|$$

con M assegnato da (7.2), e perciò

$$\|\underline{f}(\underline{x}_2)-\underline{f}(\underline{x}_1)\|\leq m\,n\,K\|\underline{x}_2-\underline{x}_1\| \qquad \forall\,\underline{x}_1,\underline{x}_2\in C.$$

Esempi

1) La funzione $f(x_1,x_2)=\dfrac{1}{x_1^2+x_2^2+1}$ è lipschitziana in R^2. Infatti le derivate parziali

$$\frac{\partial f}{\partial x_1}=\frac{-2x_1}{(x_1^2+x_2^2+1)^2}, \qquad \frac{\partial f}{\partial x_2}=\frac{-2x_2}{(x_1^2+x_2^2+1)^2}$$

esistono e sono continue in ogni punto di R^2. Inoltre

$$\max_{\underline{x}\in R^2}\left|\frac{\partial f}{\partial x_2}\right|=\max_{\underline{x}\in R^2}\left|\frac{\partial f}{\partial x_1}\right|=\max_{x_1\in R}\left|\frac{2x_1}{(x_1^2+1)^2}\right|=\frac{3\sqrt{3}}{8}.$$

Quindi f è lip $\frac{3\sqrt{3}}{4}$.

(°) Cap. I, teorema 6.2.

2) La funzione $f(x_1,x_2,x_3)=x_1x_2+x_3^2$ è lipschitziana in ogni insieme I limitato di R^3. Infatti, f è continua e dotata di derivate parziali

$$\frac{\partial f}{\partial x_1} = x_2 \quad , \quad \frac{\partial f}{\partial x_2} = x_1 \quad , \quad \frac{\partial f}{\partial x_3} = 2x_3$$

continue in R^3. Inoltre, esiste r tale che $\|\underline{x}\|<r$ per ogni $\underline{x}$ di I e nella sfera $\|\underline{x}\|<r$

$$\left|\frac{\partial f}{\partial x_i}\right| < 2r \quad , \quad i=1,2,3 .$$

Allora f è lip 6 r in tale sfera e quindi anche in I .

3) La funzione f definita da

$$f(x_1,x_2)=\begin{cases} x_1^2 e^{-\frac{x_2}{x_1}} & \text{se } x_1\neq 0 \\ 0 & \text{se } x_1=0 \end{cases}$$

è lipschitziana nel convesso

$$C=\left\{\underline{x}\in R^2 : 0\le x_2\le x_1\le 1\right\} .$$

Infatti la restrizione di f a C è continua (si noti che f non è continua nell'origine!) e dotata di derivate parziali continue in ogni punto interno a C . Inoltre internamente a C risulta

$$\left|\frac{\partial f}{\partial x_1}\right| = \left|(2x_1+x_2)e^{-\frac{x_2}{x_1}}\right| < 3$$

$$\left|\frac{\partial f}{\partial x_2}\right| = \left|-x_1 e^{-\frac{x_2}{x_1}}\right| < 1$$

e quindi f è lip 6 in C .

8. Derivate di ordine superiore.

Consideriamo una funzione $\underline{f}:R^n\to R^m$ e due versori (eventualmente coincidenti) $\underline{v}$ e $\underline{w}$ di R^n. Supponiamo che in un aperto A di R^n esista $D_{\underline{v}}\underline{f}$; essa è ovviamente una funzione di $\underline{x}$ in A . Se tale funzione è derivabile in un punto $\underline{x}°$ di A secondo il versore $\underline{w}$, chiameremo questa derivata derivata seconda di $\underline{f}$ secondo i versori $\underline{v}$ e $\underline{w}$ e la indicheremo col simbolo $D^2_{\underline{w},\underline{v}}\underline{f}(\underline{x}°)$. Quindi, per definizione

$$D^2_{\underline{w},\underline{v}}\underline{f}(\underline{x}°) = (D_{\underline{w}}D_{\underline{v}}\underline{f})(\underline{x}°)$$

qualora, ovviamente, il secondo membro esista.

Si consideri, ad esempio, la funzione

$$f(x_1,x_2)= x_1+ x_2^{4/3}$$

e siano $\underline{v}=(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}})$, $w=(\frac{1}{2},\frac{\sqrt{3}}{2})$.

f è differenziabile in ogni punto $\underline{x}$ di R^2 con

$$\text{grad } f(\underline{x}) = (1, \tfrac{4}{3}x_2^{1/3}) \quad ;$$

quindi per ogni $\underline{x}$ esiste $D_{\underline{v}}f$ ed è

$$D_{\underline{v}}f(\underline{x})=\text{grad } f(\underline{x})\cdot\underline{v}= \frac{1}{\sqrt{2}}+\frac{4}{3\sqrt{2}}x_2^{1/3} .$$

La funzione $D_{\underline{v}}f(\underline{x})$ è a sua volta differenziabile in ogni punto $\underline{x}=(x_1,x_2)$ con $x_2\neq 0$ e

$$(\text{grad } D_{\underline{v}}f)(\underline{x}) = (0, \frac{4}{9\sqrt{2}}x_2^{-2/3}) .$$

Quindi se $x_2\neq 0$ esiste $D^2_{\underline{w}\underline{v}}f$ ed è

$$D^2_{\underline{w},\underline{v}}f(\underline{x})=(\text{grad } D_{\underline{v}}f)(\underline{x})\cdot\underline{w} = \frac{2\sqrt{3}}{9\sqrt{2}}x_2^{-2/3} .$$

Se $\underline{x}=(x_1,0)$, si ha :

$$\lim_{t\to 0}\frac{D_{\underline{v}}f(\underline{x}+t\underline{w})-D_{\underline{v}}f(\underline{x})}{t} = \lim_{t\to 0}\frac{D_{\underline{v}}f(x_1+\frac{1}{2}t, \frac{\sqrt{3}}{2}t)-D_{\underline{v}}f(x_1,0)}{t}$$

$$= \lim_{t\to 0}\frac{4}{3\sqrt{2}}\left(\frac{\sqrt{3}}{2}\right)^{1/3}t^{-2/3} = \infty .$$

Quindi la derivata seconda $D^2_{\underline{w},\underline{v}}f(x_1,0)$ non esiste.

In particolare, nel caso in cui $\underline{v}$ e $\underline{w}$ siano entrambi versori di assi cartesiani, cioè $\underline{v}=\underline{e}_i$, $\underline{w}=\underline{e}_j$, la derivata seconda secondo tali versori viene chiamata derivata parziale seconda rispetto a x_i e x_j e indicata col simbolo $\frac{\partial^2 \underline{f}}{\partial x_j\partial x_i}$. Si ha cioè per definizione

$$\frac{\partial^2 \underline{f}}{\partial x_j\partial x_i}(\underline{x}) = D_{\underline{e}_j,\underline{e}_i}\underline{f}(\underline{x})=(D_{\underline{e}_j}D_{\underline{e}_i}\underline{f})(\underline{x}).$$

Ad esempio, per la funzione $f(x_1,x_2)=x_1^2x_2^3$ si ha, qualunque sia $\underline{x}$ in R^2:

$$\frac{\partial f}{\partial x_1} = 2x_1x_2^3 \ , \quad \frac{\partial f}{\partial x_2} = 3x_1^2x_2^2$$

$$\frac{\partial^2 f}{\partial x_2 \partial x_1} = \frac{\partial}{\partial x_2}\frac{\partial f}{\partial x_1} = 6x_1x_2^2$$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial}{\partial x_1}\frac{\partial f}{\partial x_2} = 6x_1x_2^2 \ .$$

In generale, l'ordine delle derivazioni <u>non</u> è invertibile, cioè in generale risulta

$$D^2_{\underline{w},\underline{v}}\underline{f}(\underline{x}) \neq D^2_{\underline{v},\underline{w}}\underline{f}(\underline{x}) \ .$$

Infatti, si consideri ad esempio la funzione $f: R^2 \to R$ definita da

$$f(x_1,x_2)= \begin{cases} \dfrac{x_1x_2^3}{x_1^2+x_2^2} & \text{se } \underline{x} \neq \underline{0} \\ 0 & \text{se } \underline{x} = \underline{0} \end{cases}$$

Se $\underline{x} \neq \underline{0}$ risulta

$$\frac{\partial f}{\partial x_1} = \frac{x_2^5 - x_1^2x_2^3}{(x_1^2+x_2^2)^2} \ , \quad \frac{\partial f}{\partial x_2} = \frac{3x_1^3x_2^2+x_1x_2^4}{(x_1^2+x_2^2)^2}$$

mentre se $\underline{x} = \underline{0}$ $\quad \dfrac{\partial f}{\partial x_1} = \dfrac{\partial f}{\partial x_2} = 0$.

Risulta quindi

$$\frac{\partial f}{\partial x_1}(0,x_2)=x_2 \quad \forall x_2 \ , \qquad \frac{\partial f}{\partial x_2}(x_1,0)=0 \quad \forall x_1 \ .$$

Di conseguenza esistono

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}(\underline{0}) = 1 \ , \quad \frac{\partial^2 f}{\partial x_1 \partial x_2}(\underline{0}) = 0$$

e non sono eguali.

Vale però il seguente

<u>Teorema</u> 8.1 - Se le due derivate $D^2_{\underline{w},\underline{v}}\underline{f}$ e $D^2_{\underline{v},\underline{w}}\underline{f}$ esistono in un intorno di $\underline{x}^\circ$ e sono continue in $\underline{x}^\circ$, allora

$$(8.1) \qquad D^2_{\underline{w},\underline{v}}\underline{f}(\underline{x}^\circ) = D^2_{\underline{v},\underline{w}}\underline{f}(\underline{x}^\circ) \ .$$

Il Teorema 9.41 p. 235 del Rudin è più preciso.

Basta dimostrare il teorema per funzioni a valori in R . Sia t reale abbastanza piccolo; poniamo

$$A(t)=f(\underline{x}^\circ+t\underline{v}+t\underline{w})-f(\underline{x}^\circ+t\underline{v})-f(\underline{x}^\circ+t\underline{w})+f(\underline{x}^\circ) \ .$$

Introduciamo la funzione ausiliaria

$$g(\tau)=f(\underline{x}^\circ+t\underline{v}+\tau\underline{w})-f(\underline{x}^\circ+\tau\underline{w}) \ .$$

Poichè t è abbastanza piccolo, g è continua e differenziabile in $[0,t]$ e risulta

$$g'(\tau)=D_{\underline{w}}f(\underline{x}^\circ+t\underline{v}+\tau\underline{w})-D_{\underline{w}}f(\underline{x}^\circ+\tau\underline{w}) \ .$$

Dal teorema dell'incremento finito si ricava allora:

$$\begin{aligned} (8.2) \qquad A&=g(t)-g(0)=g'(\vartheta t)\cdot t = \\ &=(D_{\underline{w}}f(\underline{x}^\circ+t\underline{v}+\vartheta t\underline{w})-D_{\underline{w}}f(\underline{x}^\circ+\vartheta t\underline{w}))t \end{aligned}$$

con $0<\vartheta<1$. Consideriamo ora la funzione

$$h(\tau)=D_{\underline{w}}f(\underline{x}^\circ+\tau\underline{v}+\vartheta t\underline{w}) \ .$$

Se t è abbastanza piccolo, anch'essa è continua e differenziabile in $[0,t]$ con

$$h'(\tau)=D_{\underline{v},\underline{w}}f(\underline{x}^\circ+\tau\underline{v}+\vartheta t\underline{w}) \ .$$

Poichè da (8.2) si ricava

$$A=(h(t)-h(0))t \ ,$$

applicando a h il teorema dell'incremento finito si ottiene

$$(8.3) \qquad A=h'(\vartheta' t)\ t^2=D^2_{\underline{v},\underline{w}}f(\underline{x}^\circ+\vartheta' t\underline{v}+\vartheta t\underline{w})\cdot t^2$$

con $0<\vartheta'<1$.

Ripetiamo il ragionamento, scambiando fra loro $\underline{v}$ e $\underline{w}$. Si ottiene

$$A = D^2_{\underline{w},\underline{v}}f(\underline{x}^\circ+\eta t\underline{v}+\eta' t\underline{w}) \cdot t^2$$

con $0<\eta<1$, $0<\eta'<1$. Confrontando con (8.3) si ricava allora (per $t\neq 0$)

$$D^2_{\underline{v},\underline{w}}f(\underline{x}^\circ+\vartheta' t\underline{v}+\vartheta t\underline{w})=D^2_{\underline{w},\underline{v}}f(\underline{x}^\circ+\eta t\underline{v}+\eta' t\underline{w}) \ .$$

Di qui, facendo tendere t a zero, per la continuità delle derivate in $\underline{x}^\circ$ segue la (8.1).

Con lo stesso procedimento col quale sono state definite le derivate seconde a partire dalle derivate prime, introduciamo ora la nozione di derivata di ordine $k>2$, per induzione, a partire dalle derivate di ordine $k-1$.
Siano $\underline{v}_1,\dots,\underline{v}_k$ k versori qualsiasi di R^n e la funzione $\underline{f}:R^n\to R^m$ ammetta in un intorno di $\underline{x}°$ la derivata $D^{k-1}_{\underline{v}_{k-1},\dots,\underline{v}_1}\underline{f}$ e questa derivata sia a sua volta derivabile in $\underline{x}°$ secondo il versore $\underline{v}_k$. Si dirà allora che esiste in $\underline{x}°$ la derivata k-ma $D^k_{\underline{v}_k,\dots,\underline{v}_1}\underline{f}$ e si porrà per definizione

$$D^k_{\underline{v}_k,\dots,\underline{v}_1}\underline{f}(\underline{x}°)= D_{\underline{v}_k} D^{k-1}_{\underline{v}_{k-1},\dots,\underline{v}_1}\underline{f}\,(\underline{x}°).$$

Anche per le derivate di ordine superiore al secondo vale un teorema analogo al teorema 8.1. Pertanto nel calcolo delle derivate di ordine k qualsiato si è possibile invertire l'ordine di derivazione, purchè le derivate in questione siano continue nel punto che interessa.

Ad esempio, per la funzione $f(x_1,x_2)=x_1^3x_2^{1/3}$ risulta

$$\frac{\partial^3 f}{\partial x_1^2\partial x_2}=\frac{\partial^3 f}{\partial x_1\partial x_2\partial x_1}=\frac{\partial^3 f}{\partial x_2\partial x_1^2}=2x_1x_2^{-2/3}$$

in ogni punto (x_1,x_2) con $x_2\neq 0$; se $x_2=0$ e $x_1\neq 0$ nessuna di tali derivate esiste; se infine $x_1=x_2=0$ esiste $\frac{\partial^3 f}{\partial x_2\partial x_1^2}=0$ mentre le altre due non esistono.

9. Funzioni due volte differenziabili.

La funzione $\underline{f}:R^n\to R^m$ $(n,m\geq 1)$ sia differenziabile in un aperto A. Il suo differenziale primo $d\underline{f}(\underline{x})=d\underline{f}(\underline{x};\Delta\underline{x})$ definisce, al variare di $\underline{x}$ in A e di $\Delta\underline{x}$ in R^n, una funzione a valori in R^m. Supponiamo di assegnare a $\Delta\underline{x}$ un valore fissato $\Delta\underline{x}^*$. Se, comunque sia stato scelto $\Delta\underline{x}^*\in R^n$, la funzione della variabile $\underline{x}$

(9.1) $$d\underline{f}(\underline{x};\Delta\underline{x}^*)=\sum_{i=1}^{n}\frac{\partial\underline{f}}{\partial x_i}(\underline{x})\Delta x_i^*$$

è differenziabile in un punto $\underline{x}°\in A$, diremo che $\underline{f}$ è due volte differenziabile in $\underline{x}°$.

Teorema 9.1 - $\underline{f}$ è due volte differenziabile in $\underline{x}°$, se e solo se le sue derivate parziali prime sono differenziabili in $\underline{x}°$.

Infatti, basta assumere in (9.1) $\Delta x_i^*=0$ per $i\neq j$ e $\Delta x_j^*=1$ per verificare che la condizione è necessaria $(j=1,\dots,n)$. Che la condizione sia sufficiente è ovvio.

Ad esempio, la funzione $f(x_1,x_2)=x_1x_2^2$ $(x_1,x_2\in R)$ è due volte differenziabile in ogni punto di R^2 e la funzione $f(x_1,x_2)=x_1+x_2^{4/3}$ $(x_1,x_2\in R)$ è due volte differenziabile in ogni punto (x_1,x_2) con $x_2\neq 0$ mentre non è due volte differenziabile in alcun punto dell'asse x_1.

Teorema 9.2 - Se f è due volte differenziabile in $\underline{x}°$, allora esistono in $\underline{x}°$ le derivate parziali seconde e

(9.2) $$\frac{\partial^2\underline{f}}{\partial x_i\partial x_h}(\underline{x}°)=\frac{\partial^2\underline{f}}{\partial x_h\partial x_i}(\underline{x}°)\qquad i,h=1,\dots,n.$$

Si può evidentemente supporre $m=1$. Per ogni t reale abbastanza piccolo poniamo

$$A(t)=f(\underline{x}°+t\underline{e}_i+t\underline{e}_h)-f(\underline{x}°+t\underline{e}_i)-f(\underline{x}°+t\underline{e}_h)+f(\underline{x}°).$$

Ragionando come nel teorema 8.1 si prova che

(9.3) $$A(t)=\left(\frac{\partial f}{\partial x_h}(\underline{x}°+t\underline{e}_i+\vartheta t\underline{e}_h)-\frac{\partial f}{\partial x_h}(\underline{x}°+\vartheta t\underline{e}_h)\right)t.$$

Poichè $\frac{\partial f}{\partial x_h}$ è differenziabile in $\underline{x}°$, si ha

$$\frac{\partial f}{\partial x_h}(\underline{x}°+t\underline{e}_i+\vartheta t\underline{e}_h)=\frac{\partial f}{\partial x_h}(\underline{x}°)+\operatorname{grad}\frac{\partial f}{\partial x_h}(\underline{x}°)\cdot(t\underline{e}_i+\vartheta t\underline{e}_h)+o(t)$$

$$\frac{\partial \underline{f}}{\partial x_h}(\underline{x}^\circ+\vartheta t\underline{e}_h)=\frac{\partial \underline{f}}{\partial x_h}(\underline{x}^\circ)+\operatorname{grad}\frac{\partial \underline{f}}{\partial x_h}(\underline{x}^\circ)\cdot\vartheta t\underline{e}_h+o(t)$$

per $t \to o$. Sostituendo in (9.3) si ottiene

$$A(t)=\left(\operatorname{grad}\frac{\partial f}{\partial x_h}(\underline{x}^\circ)\cdot t\underline{e}_i+o(t)\right)t =$$

$$(9.4) \qquad =t^2\,\frac{\partial^2 f}{\partial x_i\partial x_h}(\underline{x}^\circ)+o(t^2) \quad .$$

Ripetiamo il ragionamento scambiando fra loro $\underline{e}_i$ e $\underline{e}_h$; otterremo

$$A(t) = t^2\,\frac{\partial^2 f}{\partial x_h\partial x_i}(\underline{x}^\circ)+o(t^2) \quad .$$

Confrontando con la (9.4) e dividendo per t^2 si ottiene

$$\frac{\partial^2 f}{\partial x_i\partial x_h}(\underline{x}^\circ) = \frac{\partial^2 f}{\partial x_h\partial x_i}(\underline{x}^\circ) + o(1)$$

da cui, facendo tendere t a zero, si ricava la (9.2).

<u>Teorema</u> 9.3 - Se $\underline{f}$ è due volte differenziabile in $\underline{x}^\circ$, le sue derivate prime in qualsiasi direzione sono differenziabili. Inoltre esistono in $\underline{x}^\circ$ le derivate seconde di $\underline{f}$ in qualsiasi direzione e risulta

$$(9.5) \qquad D^2_{\underline{w},\underline{v}}\underline{f}(\underline{x}^\circ)=D^2_{\underline{v},\underline{w}}\underline{f}(\underline{x}^\circ)=\sum_{i,h=1}^{n}\frac{\partial^2 f}{\partial x_i\partial x_h}v_iw_h$$

qualunque siano i versori $\underline{v}$ e $\underline{w}$ di R^n.

Infatti la derivata prima

$$D_{\underline{v}}\underline{f}(\underline{x}) = \sum_{i=1}^{n}\frac{\partial \underline{f}}{\partial x_i}(\underline{x})v_i$$

è differenziabile in $\underline{x}^\circ$ per il teorema 9.1. Esiste quindi $D^2_{\underline{w},\underline{v}}\underline{f}(\underline{x}^\circ)$ e per la (3.5) si ha

$$D^2_{\underline{w},\underline{v}}\underline{f}(\underline{x}^\circ) = \Big(D_{\underline{w}}\sum_{i=1}^{n}\frac{\partial \underline{f}}{\partial x_i}v_i\Big)(\underline{x}^\circ)$$

$$= \sum_{i=1}^{n}\Big(D_{\underline{w}}\frac{\partial \underline{f}}{\partial x_i}\Big)(\underline{x}^\circ)\cdot v_i$$

$$= \sum_{i=1}^{n}\Big(\sum_{h=1}^{n}\frac{\partial^2 \underline{f}}{\partial x_h\partial x_i}(\underline{x}^\circ)w_h\Big)v_i$$

e cioè

$$(9.6) \qquad D^2_{\underline{w},\underline{v}}\underline{f}(\underline{x}^\circ)=\sum_{i,h=1}^{n}\frac{\partial^2 \underline{f}}{\partial x_h\partial x_i}(\underline{x}^\circ)v_iw_h \quad ,$$

qualunque siano i versori $\underline{v}$ e $\underline{w}$.
Allora risulta anche

$$D^2_{\underline{v},\underline{w}}\underline{f}(\underline{x}^\circ)=\sum_{i,h=1}^{n}\frac{\partial^2 f}{\partial x_i\partial x_h}(\underline{x}^\circ)v_iw_h$$

e di qui, ricordando il teorema 9.2

$$D^2_{\underline{v},\underline{w}}\underline{f}(\underline{x}^\circ)=\sum_{i,h=1}^{n}\frac{\partial^2 \underline{f}}{\partial x_h\partial x_i}(\underline{x}^\circ)v_iw_h \quad .$$

Confrontando con la (9.6) segue l'asserto.

<u>Osservazione</u> - Le ipotesi del teorema 9.3 e quelle <u>dell'analogo</u> teorema 8.1 non sono confrontabili. Infatti, l'essere una funzione due volte differenziabile in un punto non implica che due sue derivate seconde siano continue in tale punto e, viceversa, la continuità di due derivate seconde in un punto non implica che f sia ivi due volte differenziabile.
Dal teorema 5.1 si ricava immediatamente la seguente condizione sufficiente affinchè $\underline{f}$ sia due volte differenziabile in $\underline{x}^\circ$.

<u>Teorema</u> 9.4 - In un intorno di $\underline{x}^\circ$ esistano e siano continue le derivate parziali prime di $\underline{f}$. Inoltre esistano in questo intorno anche le derivate parziali seconde e siano continue in $\underline{x}^\circ$. Allora $\underline{f}$ è due volte differenziabile in $\underline{x}^\circ$.

<u>Corollario</u> 9.1 - Se in un aperto A $\underline{f}$ è dotata di derivate parziali prime e seconde continue, allora $\underline{f}$ è due volte differenziabile in ogni punto di A .
Una tale funzione viene detta <u>due volte differenziabile con continuità in</u> A, oppure <u>di classe</u> $\mathscr{C}^{(2)}(A)$.

Introduciamo ora la nozione di differenziale secondo. Se $\underline{f}$ è due volte differenziabile in $\underline{x}^\circ$ chiameremo <u>differenziale secondo</u> di $\underline{f}$ in $\underline{x}^\circ$ il differenziale nel punto $\underline{x}^\circ$ del suo differenziale primo, cioè

$$d^2\underline{f}(\underline{x}^\circ) = (d(d\underline{f}))(\underline{x}^\circ) \quad .$$

Teorema 9.5 - Se $\underline{f}$ è due volte differenziabile in $\underline{x}^\circ$ risulta

$$(9.7) \quad d^2\underline{f}(\underline{x}^\circ)=d^2\underline{f}(\underline{x}^\circ;\Delta\underline{x}^*,\Delta\underline{x})=\sum_{i,h=1}^{n} \frac{\partial^2 \underline{f}}{\partial x_h \partial x_i}\, \Delta x_h \Delta x_i^* \quad .$$

Infatti, da 9.1 si ricava

$$d^2\underline{f}(\underline{x}^\circ)=d\Big(\sum_{i=1}^{n} \frac{\partial \underline{f}}{\partial x_i}\, \Delta x_i^*\Big)(\underline{x}^\circ)=\sum_{i=1}^{n} \Big(d\, \frac{\partial \underline{f}}{\partial x_i}\Big)(\underline{x}^\circ)\Delta x_i^*$$
$$=\sum_{i=1}^{n} \Big(\sum_{h=1}^{n} \frac{\partial^2 \underline{f}}{\partial x_h \partial x_i}\, \Delta x_h\Big)\, \Delta x_i^*$$

e cioè la (9.7).

simmetrica

Osservazione - Il differenziale secondo $d^2\underline{f}(\underline{x}^\circ)$ per ogni $\Delta\underline{x}^*$ fissato è lineare in $\Delta\underline{x}$ e viceversa, cioè è una forma bilineare nelle due variabili $\Delta\underline{x}^*$ e $\Delta\underline{x}$ definita in $R^n \times R^n$ a valori in R^m.

10. Funzioni k volte differenziabili (k > 2).

Iterando il procedimento seguito per definire le funzioni due volte differenziabili e il differenziale secondo, introduciamo ora il concetto di funzione k volte differenziabile (k > 2) e di differenziale k-mo.
La funzione $\underline{f}:R^n \to R^m$ $(n,m \geq 1)$ sia k-1 volte differenziabile in un aperto $A \subset R^n$. Se il suo differenziale di ordine k-1

$$d^{k-1}\underline{f}(\underline{x})=d^{k-1}\underline{f}(\underline{x};\Delta\underline{x}^{(1)},\ldots,\Delta\underline{x}^{(k-1)}) \,,$$

comunque siano fissati $\Delta\underline{x}^{(1)},\ldots,\Delta\underline{x}^{(k-1)}$ in R^n, è differenziabile in un punto $\underline{x}^\circ \in A$, diremo che $\underline{f}$ è k volte differenziabile in $\underline{x}^\circ$ e chiameremo differenziale k-mo di $\underline{f}$ in $\underline{x}^\circ$ $d^k\underline{f}(\underline{x}^\circ)$ il differenziale primo in $\underline{x}^\circ$ del suo differenziale di ordine k-1.
Per le funzioni k volte differenziabili valgono risultati analoghi a quelli ottenuti nel par.9 per le funzioni due volte differenziabili.

Segnaliamo in particolare i seguenti:

1) $\underline{f}$ è k volte differenziabile in $\underline{x}^\circ$ se e solo se:
a) esistono continue in un intorno di $\underline{x}^\circ$ le derivate parziali fino all'ordine k-2 incluso;
b) esistono in tale intorno le derivate parziali di ordine k-1 e sono differenziabili in $\underline{x}^\circ$.

2) Se $\underline{f}$ è k volte differenziabile in $\underline{x}^\circ$, esistono in tale punto le derivate di ordine k, l'ordine di derivazione è invertibile e vale la formula

$$D_{\underline{v}^{(1)},\ldots,\underline{v}^{(k)}}\underline{f}(\underline{x}^\circ)=\sum_{i_1,\ldots,i_k=1}^{n} \frac{\partial^k \underline{f}}{\partial x_{i_1}\cdots\partial x_{i_k}}\, v_{i_1}^{(1)}\cdots v_{i_k}^{(k)}$$

qualunque siano i versori $\underline{v}^{(1)},\ldots,\underline{v}^{(k)}$ di R^n.

3) Se in un intorno di $\underline{x}^\circ$ esistono e sono continue tutte le derivate parziali di $\underline{f}$ fino all'ordine k-1 incluso e inoltre esistono in tale intorno le derivate parziali di ordine k e sono continue in $\underline{x}^\circ$, allora $\underline{f}$ è k volte differenziabile in $\underline{x}^\circ$.

4) Se in un aperto A esistono e sono continue tutte le derivate parziali di $\underline{f}$ fino all'ordine k incluso, allora $\underline{f}$ è k volte differenziabile in A. Una tale funzione viene detta k volte differenziabile con continuità in A, oppure di classe $\mathcal{C}^{(k)}(A)$.

5) Se $\underline{f}$ è k volte differenziabile in $\underline{x}^\circ$, risulta

$$(10.1) \quad d^k\underline{f}(\underline{x}^\circ)=d^k\underline{f}(\underline{x}^\circ;\Delta\underline{x}^{(1)},\ldots,\Delta\underline{x}^{(k)})=$$
$$=\sum_{i_1,\ldots,i_k=1}^{n} \frac{\partial^k \underline{f}}{\partial x_{i_1}\cdots\partial x_{i_k}}(\underline{x}^\circ)\, \Delta x_{i_1}^{(1)}\cdots \Delta x_{i_k}^{(k)}.$$

Il differenziale k-mo in $\underline{x}^\circ$ è quindi una forma k-lineare a valori in R^m definita al variare di $\Delta\underline{x}^{(1)},\ldots,\Delta\underline{x}^{(k)}$ in R^n.
Denotiamo ora per semplicità di scrittura, con $d^k\underline{f}(\underline{x}^\circ;\Delta\underline{x})$ la funzione dei punti di R^n a valori in R^m definita da

$$(10.2) \quad d^k\underline{f}(\underline{x}^\circ;\Delta\underline{x})=d^k\underline{f}(\underline{x}^\circ;\Delta\underline{x},\ldots,\Delta\underline{x})=$$

$$=\sum_{i_1,\dots,i_k=1}^{n} \frac{\partial^k \underline{f}}{\partial x_{i_1}\cdots\partial x_{i_k}}(\underline{x}^\circ)\Delta x_{i_1}\cdots\Delta x_{i_k}.$$

La funzione di $\Delta\underline{x}$ così definita è un polinomio omogeneo di grado k nelle componenti di $\Delta\underline{x}$. Tenuto conto dell'invertibilità dell'ordine di derivazione, tale polinomio può essere scritto come potenza formale nel modo seguente:

$$d^k\underline{f}(\underline{x}^\circ;\Delta\underline{x})=\Big(\sum_{i=1}^{n}\frac{\partial}{\partial x_i}\Delta x_i\Big)^{(k)}\underline{f}(\underline{x}^\circ) \quad .$$

Osserviamo anche che per $\Delta\underline{x}\to\underline{0}$ risulta

$$d^k\underline{f}(\underline{x}^\circ;\Delta\underline{x}) = O(\|\Delta\underline{x}\|^k) \quad .$$

Poichè nel seguito ci accadrà di considerare il differenziale di ordine k (10.1) <u>esclusivamente</u> per $\Delta\underline{x}^1=\dots=\Delta\underline{x}^k=\Delta\underline{x}$, <u>conveniamo</u>, con un abuso evidente (che non può però dare luogo ad equivoci nel seguito), di chiamare <u>differenziale k-mo</u> di $\underline{f}$ in $\underline{x}^\circ$ anche la funzione definita in (10.2) e di denotare tale funzione col simbolo $d^k\underline{f}(\underline{x}^\circ)$, quando ciò non dia luogo a inconvenienti.

(vd. Cartan, p. 78)

11. <u>Formula di Taylor. Resto nella forma di Peano</u>.

Per le funzioni k volte differenziabili in un punto vale il seguente teorema, che estende alle funzioni di più variabili la formula di Taylor col resto nella forma di Peano.

<u>Teorema</u> 11.1 - Se $\underline{f}:R^n\to R^m$ $(n,m\geq 1)$ è k volte <u>differenziabile</u> nel punto $\underline{x}^\circ$, allora per $\underline{h}\to\underline{0}$ risulta

$$(11.1)\qquad \underline{f}(\underline{x}^\circ+\underline{h})=\underline{f}(\underline{x}^\circ)+\frac{d\underline{f}(\underline{x}^\circ;\underline{h})}{1!}+\dots+\frac{d^k\underline{f}(\underline{x}^\circ,\underline{h})}{k!}+ o(\|\underline{h}\|^k) \quad .$$

Per dimostrare il teorema, cominciamo a osservare che per k=1 la (11.1) è senz'altro vera perchè coincide con la definizione di funzione differenziabile in $\underline{x}^\circ$. Procederemo quindi per induzione supponendo che la (11.1) valga per k-1 e dimostrandola per k. Potremo inoltre evidentemente supporre m=1.

Poniamo, per semplicità di scrittura,

$$g(\underline{h})=f(\underline{x}^\circ+\underline{h})-f(\underline{x}^\circ)-\sum_{j=1}^{k}\frac{1}{j!}d^jf(\underline{x}^\circ) \quad .$$

Per dimostrare il teorema basterà provare che

$$(11.2)\qquad g(\underline{h}) = o(\|\underline{h}\|^k) \quad \text{per} \quad \underline{h}\to\underline{0} \quad .$$

A tale scopo, cominciamo con l'osservare che, poichè f è k volte differenziabile in $\underline{x}^\circ$, f è differenziabile almeno una volta (essendo $k\geq 2$) in un intorno di $\underline{x}^\circ$. Quindi, se $\|\underline{h}\|$ è abbastanza piccola, g è differenziabile e risulta

$$\frac{\partial g}{\partial h_i}(\underline{h})=\frac{\partial f}{\partial x_i}(\underline{x}^\circ+\underline{h})-\sum_{j=1}^{k}\frac{1}{j!}\frac{\partial}{\partial h_i}d^jf(\underline{x}^\circ) \quad .$$

Con alcuni calcoli (elementari, benchè non tutti semplici!) si può verificare che

$$\frac{\partial}{\partial h_i}d\,f(\underline{x}^\circ) = \frac{\partial f}{\partial x_i}(\underline{x}^\circ)$$

e che, se $j=2,\dots,k$

$$\frac{\partial}{\partial h_i}d^jf(\underline{x}^\circ)=\frac{\partial}{\partial h_i}\Big(\sum_{\ell=1}^{n}\frac{\partial}{\partial x_\ell}h_\ell\Big)^{(j)}f(\underline{x}^\circ) = j\Big(\sum_{\ell=1}^{n}\frac{\partial}{\partial x_\ell}h_\ell\Big)^{(j-1)}\frac{\partial f}{\partial x_i}(\underline{x}^\circ) \quad .$$

Di conseguenza, se $\|\underline{h}\|$ è abbastanza piccola, risulta

$$\frac{\partial g}{\partial h_i}(\underline{h})=\frac{\partial f}{\partial x_i}(\underline{x}^\circ+\underline{h})-\frac{\partial f}{\partial x_i}(\underline{x}^\circ)-\sum_{j=1}^{k-1}\frac{1}{j!}\Big(d^j\frac{\partial f}{\partial x_i}\Big)(\underline{x}^\circ).$$

Dalla (11.1) (valida per k-1) applicata alla funzione $\frac{\partial f}{\partial x_i}$ si ricava

$$\frac{\partial g}{\partial h_i}(\underline{h}) = o(\|\underline{h}\|^{k-1}) \quad \text{per} \quad \underline{h}\to\underline{0} \quad .$$

Dal teorema dell'incremento finito (7.1) applicato a $g(\underline{h})-g(\underline{0})=g(\underline{h})$ si ricava allora

$$|g(\underline{h})|=o(\|\underline{h}\|^{k-1})\cdot\|\underline{h}\| = o(\|\underline{h}\|^k)$$

e cioè vale (11.2), c.d.d.

E s e m p i

1) La funzione $f(x_1,x_2)=(x_1+x_2)\text{sen}\, x_2$ è di classe $\mathcal{C}^\infty$ in R^2 e quindi è rappresentabile nella forma (11.1) nell'intorno di qualsiasi punto $\underline{x}^\circ \in R^2$, qualunque sia k.
Determiniamo per esempio la formula di Taylor (11.1) per questa funzione nel caso k=3, rispettivamente nei punti $(0,\pi/2)$ e $(0,0)$.
Qualunque sia il punto $(x_1,x_2)\in R^2$ risulta

$$\frac{\partial f}{\partial x_1}=\text{sen}\, x_2 \;;\quad \frac{\partial f}{\partial x_2}=(x_1+x_2)\cos x_2+\text{sen}\, x_2 \;;$$

$$\frac{\partial^2 f}{\partial x_1^2}=0 \;;\quad \frac{\partial^2 f}{\partial x_1\partial x_2}=\cos x_2 \;;\quad \frac{\partial^2 f}{\partial x_2^2}=-(x_1+x_2)\text{sen}\, x_2+2\cos x_2 \;;$$

$$\frac{\partial^3 f}{\partial x_1^3}=\frac{\partial^3 f}{\partial x_1^2\partial x_2}=0 \;;\quad \frac{\partial^3 f}{\partial x_1\partial x_2^2}=-\text{sen}\, x_2 \;;$$

$$\frac{\partial^3 f}{\partial x_2^3}=-(x_1+x_2)\cos x_2-3\,\text{sen}\, x_2 \;.$$

Di conseguenza si ha

$$\left(\frac{\pi}{2}+h_1+h_2\right)\text{sen}\left(\frac{\pi}{2}+h_2\right)=f\left(0+h_1,\frac{\pi}{2}+h_2\right)=\frac{\pi}{2}+h_1+h_2-\frac{\pi}{4}h_2^2-\frac{1}{2}h_2^3+$$
$$+\,o((h_1^2+h_2^2)^{3/2}) \;;$$

$$(x_1+x_2)\text{sen}\, x_2=f(0+x_1,0+x_2)=x_1x_2+x_2^2+o((x_1^2+x_2^2)^{3/2}) \;.$$

2) La funzione $f(x_1,x_2)=x_1^2+x_2e^{x_1}$ è di classe $\mathcal{C}^\infty$ in R^2 ed è quindi rappresentabile nella forma (11.1) nell'intorno di qualsiasi punto di R^2, qualunque sia k. La relativa formula (11.1) può essere facilmente ottenuta nel modo seguente.
Consideriamo p.e. il punto $\underline{x}^\circ=(2,1)$. Posto $x_1=2+h_1$, $x_2=1+h_2$ si ha

$$f(2+h_1,1+h_2)=(2+h_1)^2+(1+h_2)e^{2+h_1}=4+4h_1+h_1^2+e^2(1+h_2)e^{h_1} \;.$$

Poichè la funzione $g(h_1)=e^{h_1}$ ammette in un intorno di $h_1=0$ uno sviluppo di Taylor della forma

$$e^{h_1}=\sum_{n=0}^{k}\frac{1}{n!}h_1^n+o(h_1^k)$$

si ottiene

$$f(2+h_1,1+h_2)=4+4h_1+h_1^2+e^2\sum_{n=0}^{k}\frac{1}{n!}h_1^n+e^2\sum_{n=0}^{k}\frac{1}{n!}h_1^n h_2+(1+h_2)o(h_1^k)=$$
$$=4+e^2+(4+e^2)h_1+e^2h_2+\left(1+\frac{1}{2}e^2\right)h_1^2+e^2h_1h_2+\sum_{n=3}^{k}\frac{1}{n!}e^2h_1^3+$$

$$+\sum_{n=2}^{k-1}\frac{1}{n!}e^2h_1^nh_2+o((h_1^2+h_2^2)^{k/2}) \;.$$

3) La funzione $f(x_1,x_2)=(x_1-x_2)(x_1+x_2^2)$ è rappresentabile nella forma (11.1) in un intorno dell'origine qualunque sia k. Precisamente, essendo

$$f(x_1,x_2)=x_1^2-x_1x_2+x_1x_2^2-x_2^3$$

la relativa formula (11.1) diviene

$$f(x_1,x_2)=o((x_1^2+x_2^2)^{1/2}) \qquad \text{se} \quad k=1 \;;$$
$$f(x_1,x_2)=x_1^2-x_1x_2+o((x_1^2+x_2^2)) \qquad \text{se} \quad k=2$$

mentre se $k\geq 3$ si ha

$$f(x_1x_2)=x_1^2-x_1x_2+x_1x_2^2-x_3^2$$

ed il termine complementare è identicamente nullo.

4) La funzione $f(x_1,x_2)=(e^{x_1^2}+x_2^2-1)^{3/2}$ è di classe $\mathcal{C}^\infty$ in $R^2-\{\underline{0}\}$ e quindi è rappresentabile nella forma (11.1) nell'intorno di qualsiasi punto di R^2 diverso dall'origine, qualunque sia k.
Inoltre se $k\leq 2$ essa è rappresentabile nella forma (11.1) anche in un intorno dell'origine. Infatti

$$e^{x_1^2}+x_2^2-1=1+x_1^2+o(x_1^4)+x_2^2-1=x_1^2+x_2^2+o(x_1^4)=$$
$$=\|\underline{x}\|^2+o(\|x\|^2)=\|x\|^2(1+o(1))$$

da cui

$$(e^{x_1^2}+x_2^2-1)^{3/2}=\|x\|^3(1+o(1))^{3/2}=\|x\|^3+o(\|x\|^3) \;.$$

Quindi f è due volte differenziabile nell'origine con $df=d^2f=0$ e vale (11.1) per k=1 e k=2.
Se $k>2$ la (11.1) non vale perchè $\|x\|^3$ (e di conseguenza f) non è tre volte differenziabile nell'origine.

12. Formula di Taylor. Resto nella forma di Lagrange.

Per le funzioni reali di più variabili k volte differenziabili in un aperto A vale un teorema che estende a tali funzioni la formula di Taylor col resto nella forma di Lagrange. Esso costituisce anche una generalizzazione del teorema dell'incremento finito 6.1.

Questa dimostrazione è essenzialmente unidimensionale.
Allo stesso modo si dimostra la formula di Taylor con resto in forma integrale.

Teorema 12.1 - La funzione $f: R^n \to R$ sia k volte differenziabile in un aperto A, e il segmento chiuso di estremi $\underline{x}^o$, $\underline{x}^o+\underline{h}$ appartenga ad A. Allora esiste $\vartheta = \vartheta(\underline{x}^o, \underline{h})$, $0 < \vartheta < 1$, tale che

$$(12.1) \quad f(\underline{x}^o+\underline{h}) = f(\underline{x}^o) + \frac{df(\underline{x}^o;\underline{h})}{1!} + \dots + \frac{d^{k-1}f(\underline{x}^o;\underline{h})}{(k-1)!} + \frac{d^k f(\underline{x}^o+\vartheta\underline{h};\underline{h})}{k!} \quad .$$

Infatti, se $n=1$, la (12.1) equivale a

$$(12.2) \quad f(x^o+h) = f(x^o) + \sum_{j=1}^{k-1} \frac{1}{j!} \frac{d^j f}{dx^j}(x^o)h^j + \frac{1}{k!}\frac{d^k f}{dx^k}(x^o+\vartheta h)h^k$$

con ϑ opportuno, $0 < \vartheta < 1$, formula ben nota.

Se $n > 1$, consideriamo la funzione ausiliaria $g: R \to R$ $g(t) = f(\underline{x}^o+t\underline{h})$. Essa è composta della funzione $\varphi: R \to R^n$ $\varphi(t) = \underline{x}^o+t\underline{h}$, di classe $\mathscr{C}^\infty(R)$, e della funzione f ed è quindi k volte differenziabile con continuità in $[0,1]$. Da (12.2) si ricava allora

$$(12.3) \quad g(1)-g(0) = \sum_{j=1}^{k-1} \frac{1}{j!}\frac{d^j g}{dt^j}(0) + \frac{1}{k!}\frac{d^k g}{dt^k}(\vartheta)$$

con ϑ opportuno, $0 < \vartheta < 1$.

Ma

$$(12.4) \quad g(1)-g(0) = f(\underline{x}^o+\underline{h}) - f(\underline{x}^o) \quad .$$

Inoltre, per ogni $t \in [0,1]$ si ha

$$g'(t) = \text{grad } f(\underline{x}^o+t\underline{h})\cdot\underline{h} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\underline{x}^o+t\underline{h})\cdot h_i = df(\underline{x}^o+t\underline{h}) \quad ;$$

$$g''(t) = \frac{d}{dt}\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(\underline{x}^o+t\underline{h})h_i = \sum_{i=1}^{n}\left\{\text{grad }\frac{\partial f}{\partial x_i}(\underline{x}^o+t\underline{h})\cdot\underline{h}\right\}h_i = \sum_{i,j=1}^{n}\frac{\partial^2 f}{\partial x_i \partial x_j}(\underline{x}^o+t\underline{h})h_i h_j = d^2 f(\underline{x}^o+t\underline{h})$$

. . . . .

Vale anche per $m>1$, con il resto opportuno.

e così via. Si ha quindi, per ogni $t \in [0,1]$ e per ogni $j=1,2,\dots,k$

$$(12.5) \qquad g^{(j)}(t) = d^j f(\underline{x}^o+t\underline{h}) \quad .$$

Da (12.3), (12.4) e (12.5) si ricava allora la (12.1).

<u>Osservazioni</u> - 1) Se $k=1$, il teorema 12.1 si riduce al teorema dell'incremento finito 6.1.

2) La formula (12.1) non è in generale valida per funzioni $\underline{f}: R^n \to R^m$ se $m>1$. (Si veda il contro-esempio all'inizio del paragrafo 7).

3) Il teorema 12.1 vale anche con le meno restrittive ipotesi seguenti:

a) f sia definita in un insieme contenente il segmento chiuso S di estremi $\underline{x}^o$, $\underline{x}^o+\underline{h}$ e la restrizione di f a S sia continua;
b) f sia k-1 volte differenziabile in $\underline{x}^o$;
c) f sia k volte differenziabile nei punti interni ad S;
d) le restrizioni a $S-\{\underline{x}^o+\underline{h}\}$ delle derivate parziali di f fino all'ordine k-1 incluso siano continue.

La relazione

$$g^{(j)}(t) = f^{(j)}(x_o+th; h,\dots,h) \qquad j=0,1,2,\dots,k$$

va dimostrata per induzione.

$j=0$ ovvio

vero $j \Rightarrow$ vero $j+1$

$$g^{(j+1)}(t) = \frac{d}{dt}\left(f^{(j)}(x_o+th; h,\dots,h)\right)$$

$$= \left[f^{(j)}(\cdot\,; h,\dots,h)\right]'(x_o+th)\cdot h$$

$$= f^{(j+1)}(x_o+th; \underbrace{h,\dots,h}_{j+1 \text{ termini}})$$